

Capacity and performance analysis for multi-user system under distributed opportunistic scheduling in a time dependent channel

O. Shmuel, A. Cohen, and O. Gurewitz

Abstract

Consider the problem of a multiple access channel with a large number of users. While several multi-user coding techniques exist, in practical scenarios, not all users can be scheduled simultaneously. Thus, a key problem is which users to schedule in a given time slot. Although the problem has been studied in various time-independent scenarios, capacity scaling laws and algorithms for time-dependent channels (e.g., Markov channels) remains relatively unexplored. The basic assumption that users' capacities are *i.i.d.* random variables is no longer valid, and a more realistic approach should be taken. The channel observed by a user is time varying, hence its distribution changes in time, i.e., a channel with memory. Since intelligent user selection has significant advantages, analyzing the distribution of the maximal capacity seen by a user in the system can give a good approximations for the capacity scaling law.

In this work, we consider a channel with two channel states, Good and Bad, where in each point in time a user is associated with one of them. The process of moving between states is modeled by a Markovian process (Gilbert-Elliott Channel). First, we derive the expected capacity under scheduling for this time dependent environment and show that its scaling law is $O(\sigma_g \sqrt{2 \log K} + \mu_g)$, where σ_g, μ_g are the good channel parameters (assuming Gaussian capacity approximation, e.g., under MIMO) and K is the number of users. Analysis uses Extreme Value Theory (EVT) along with finding the suitable normalizing constants.

In addition, a distributed algorithm for this scenario is suggested along with its channel capacity analysis. The algorithm is threshold-based and the rate for exceeding it is analyzed using Point Process Approximation (PPA). The analysis is done using the properties of the extremes of chain dependent sequences, which is proved in this study to converge to one of the extreme value distributions. The expected capacity, while imposing the discussed algorithm, scales as $O(e^{-1}(\sigma_g \sqrt{2 \log K} + \mu_g))$, hence, there is no loss in optimality due to the distributed algorithm. The foundation for extending this study to more channel states is set and can be easily done.

Finally, we turn to performance analysis of such systems while assuming the users are not necessarily fully backlogged, focusing on the queueing problem and, especially, on the strong dependence between the queues. We adopt the celebrated model of Ephremides and Zhu to give new results on the convergence of the probability of collision to its average value (as the number of users grows), and hence for the ensuing system performance metrics, such as throughput and delay. We further utilize this finding to suggest a much simpler approximate model, which accurately describes the system behavior when the number of queues is large. The system performance as predicted by the approximate models shows excellent agreement with simulation results.

Contents

Abstract	I
1 Introduction	2
1.1 Main Contributions	3
2 Related Work	4
2.1 Scheduling Algorithms	4
2.1.1 Broadcast Channel	4
2.1.2 Multiple-Access Channel	6
2.2 Time Dependent Channel	7
2.3 Limit laws in time dependent channels	7
2.4 Queueing system	8
3 Preliminaries	9
3.1 The Gaussian Channel	9
3.2 Capacity of MIMO	10
3.2.1 multi-antenna Gaussian channels	10
3.2.2 A Gaussian Approximation to the MIMO Capacity	11
3.3 Extreme Value Theory	11
3.4 Point Process Approximation	14
3.5 Chain Dependent Process	15
4 Model Description	17
4.1 Time dependent channel	17
5 Capacity under time dependent channel	19
5.1 Capacity - Centralized Scheduling	19
5.2 Capacity - Distributed Scheduling	22
5.2.1 Distributed Algorithm	25
5.2.2 Threshold Estimation	26
5.2.3 Refinement On The Threshold Value	27
6 Performance analysis - distributed algorithm	28
6.1 Queueing Approximate model I	29
6.1.1 Review on the approximate model	29
6.1.2 Performance analysis and comparisons	32
6.2 Queueing Approximate model II	34
6.2.1 Threshold exceedance process	34
6.2.2 Service time analysis	37
6.3 Queueing Approximate model III	40
7 Conclusions	44

A	Convergence to Extreme Value Distribution	45
A.1	Convergence proof	45
A.2	Normalizing constants	48
B	Capacity - Direct expression analysis	50
B.1	Capacity - analyzing change in group sizes	50
B.1.1	Transition probability matrix	50
B.1.2	Stationary distribution	52
B.1.3	Capacity analyze	53
B.1.4	Capacity Bounds	55
C	Random selection Proof	58
D	System-status chain transition probabilities	59
E	Average success probabilities	62

Chapter 1

Introduction

Numerous innovations, such as sophisticated coding techniques and advances in the endpoint equipment, have been suggested over recent years in order to cope with the high traffic demands over the last-hop wireless medium. Yet, the rapid growth in the number of users and devices which are connected via this last mile wireless access architecture requires efficient scheduling techniques as well, since in practical scenarios, not all of them can be scheduled together. Indeed, scheduling schemes have been widely studied in such multiuser environments, trying to exploit the multiuser diversity which is inherent in wireless medium, by scheduling users when their instantaneous channel quality is at its best. Utilizing multiple antennas both at the transmitter and the receiver has also been adopted by state of the art wireless standards such as 802.11n, 802.11ac and LTE advanced. Multiuser diversity strategies under such Multiple-Input Multiple-Output (MIMO) technologies have been exploited in many scheduling studies, e.g., in [1], [2], where Zero-Forcing Beamforming (ZFBF) was investigated and user selection was performed in order to avoid interference among users streams.

Scheduling schemes can be categorized as *centralized*, typically requiring high overhead in collecting channel reports from many users by the central entity (e.g., access point or base station), in order to perform informed user selection, or *distributed*, which are designed such that the users can schedule themselves. One such opportunistic distributed scheme which has shown excellent exploitation of multiuser diversity, lets the users transmit based on their instantaneous channel condition. Specifically, users may transmit only if their channel gain is above a given threshold.

While the problem in its many forms, as described above, has been studied in various time-independent scenarios, capacity and algorithms for time-dependent channels (e.g., Markov channels) remained relatively unexplored. We consider the situation for the upstream traffic in a distributed manner with the assumption that a user experiences a time varying channel, therefore a time dependent model is suggested, which reflects the time varying channel distribution. Using this model and its properties we were able to achieve closed analytic expressions in different ways for the channel capacity while using two statistical tools, Extreme Value Theory (EVT) and Point Process Approximation (PPA). Those enabled us to examine the limit distribution of the maximal capacity value for the transmitting user, i.e. multi-user diversity. We also looked on a model that describes the segmentation of users according to the different channel states they exist in, and tried to obtain also the expression for the channel capacity.

Obviously, a distributive paradigm in which users schedule themselves for transmission can result in collisions between users, which, in turn, affects channel utilization and quality of service (QoS). In this work we are interested in characterizing these important metrics. We study users' performance under realistic traffic patterns, in which users are not fully backlogged and the arrival process of each user is a random process. Specifically, we assume a realistic model in which the

arrival process of each user is modeled as a Poisson process. Each user transmits packets in a first-in-first-out (FIFO) manner, such that packets arriving to a user which is already busy with a pending packet transmission, wait for their turn to be transmitted. Accordingly, each user maintains a queue in which it stores packets waiting for transmission. We study performance metrics other than capacity, such as delay and buffer occupancy. We suggest a general model, consisting with a system of K users in a random, multiple-access system who interact and affect each other. This theoretical problem was studied in the past and it has been shown that the interaction between the queues in such systems makes the analysis inherently difficult. Thus there is no clear understanding of the performance of such systems with large number of users (more than 3), due to the complexity in describing the strong interdependence between the user's queues. Only approximate models were developed in order to describe the system behaviour.

This work is organized as follows. In Chapter 2, literature survey is given which will review related work in the subject. In Chapter 3 we will review some of the relevant materials for better understanding the problem and its properties. When part of which will cover some results concerning EVT and PPA. Those are necessary to carry out the asymptotic analysis of the channel capacity. We will elaborate on results that are relative for time dependent processes (stationary sequences), which we shall use in our time dependent situation. In Chapter 4 we will describe our model of the system and basic assumptions for this work. As mentioned before, this work is divided to two principal areas of research, the first is given in chapter 5, there we will focus on the design and analysis of an algorithm which schedules users in the time-dependent regime in a distributed manner. The algorithm is threshold based, and takes the advantages of Multi-User diversity in its operation. The resulting channel capacity is given analytically, simulations and bounds are also given. In chapter 6, the second area will be given, which will examine the performance analysis of the queues in our distributed multiple-access system. Models for the independent and dependent time scenarios is presented along with numerical results.

1.1 Main Contributions

- In chapter 5, we start with giving a closed expression for the channel capacity scaling laws, in a centralized time-dependent environment, while the best user, with the highest channel quality, is chosen to utilize the channel. Then we present a distributed algorithm for the users scheduling mechanism, and the time dependent channel capacity scaling laws under this algorithm. During the analysis of the distributed algorithm, we show a proof for the extreme values convergence of our chain dependent process to one of the EVT distributions, while using the dependency condition D and D' for stationary sequences. And as part of this convergence, we derive the normalizing constants for the convergence to extreme value distribution, where the marginal distribution is a mixture of normal distributions. All of these analytical developments, set foundations for the extension of the channel model to more than two states, and finding the maximal value distribution for this case to derive the expected channel capacity.
- In chapter 6, we present approximate models for the systems' queues behaviour. We first concentrate and review the model presented in [3], while fitting it to our setting. Numerical results and interesting observations are presented for this elaborated model. Then, we present a simpler approximate model which in contrary to the former, is able to describe the systems' behaviour when the number of queues is large. Lastly, we suggest a time dependent queue model which relates to our time dependent channel and shows very good agreement with simulations results.

Chapter 2

Related Work

Two of the most important results in network information theory are the capacity of the Gaussian multiple-access channel (MAC) and the capacity of the Gaussian broadcast channel (BC). The desire to maximally utilize both channel capacities is in the heart of many of the studies dealing with shared medium communication. As mentioned in the pervious chapter, scheduling and user selection is a crucial part in achieving these capacities when it comes to large population. In this study we are dealing with the Gaussian MAC model, nevertheless many ideas from works concerning the Gaussian BC are extremely relevant such as the utilization of multi-user diversity and the use of order statistic (EVT) for statistical deductions. In the following sections we review some of state of the art works concerning both types of channels and the attempts to obtain the optimal capacity by using different methods for scheduling and user selection along with variety of statistical analysis.

The communication system considered in this work is single user MIMO - Multiple Input Multiple Output. Basically it means to use multiple antennas at both the transmitter and the receiver in order to increase data rates through multiplexing or to improve performance through diversity. Multiplexing is obtained by exploiting the structure of the channel gain matrix to obtain independent signalling paths that can be used to send independent data.

By using this system we can achieve high data rate without increasing the total transmission power or bandwidth. Additionally the use of multiple antennas at both the transmitter and receiver provides significant increase in the capacity [4–7]. MIMO has been supported by several novel wireless standards such as 802.11ac and LTE Advanced.

2.1 Scheduling Algorithms

While considering a large multi-user networks, each user will experience different channel state which caused by various conditions. This diversity in users' channel state may be used to enhanced the ergodic capacity of the channel. In [8], it was shown that in order to improve spectral efficiency, at any given time, only the user (or users) with the strongest signal should transmit over the entire bandwidth. Hence, we must choose intelligently the user with minimum complexity and overhead. The work of [9] improved the work of [8] when considering delay aspects, when they treat the situation of starvation. Another work which strengthened the fact that appropriate user selection is essential, and in several cases can even achieve optimality with sub-optimal detectors is [10], which discussed various user selection methods in several MIMO detection schemes.

2.1.1 Broadcast Channel

Dirty Paper Coding (DPC) for Gaussian broadcast channel, have proven to be the optimal solution for achieving the channel capacity [11]. In practice this technique is not used as it involve complex

computations and it is not feasible for large population of users. Hence, with the use of multi-user diversity, simpler methods for achieving channel capacity were derived. In [2], the researchers investigate a technique to achieve the MU-MIMO capacity called Zero-Forcing Beamforming (ZFBF). The paper investigates if, for a large number of users, simpler schemes can achieve the same performance as dirty paper coding (DPC) technique.

Beamforming (BF) is a strategy that can serve multiple users at a time, where each user stream is coded independently and multiplied by a beamforming weight vector for transmission through multiple antennas. By choosing the weight vectors mutual interference between different streams could be reduced or disappear, allowing spatial separation between users and as a result support multiple users simultaneously. In the paper the researchers have managed to present a low complexity algorithm for semi orthogonal user selection, which achieves a sum rate close to the optimal rate given by DPC. In addition they presented a fairly scheduling scheme that is asymptotically optimal and has good performance.

As BF and DPC show similar performances of the sum rate capacity, a comparison was made with time sharing capacity in [12], and the researchers' conclusion was that the sum rate of DPC and BF outperform of the time sharing if the transmitter is equipped with multiple antennas. In [13], asymptotically optimal scheme for downstream traffic for coordinated beamforming where multiple BSs jointly optimize their respective beamforming vectors was given. In [1], another scheduling algorithm over a MIMO broadcast channel is introduced. Given a set of users, the scheduler selects more than one user and transmits independent data to them simultaneously by using zero-forcing beamforming. Taking computational complexity into account, a greedy method finding the best and most orthogonal channel vectors is proposed. Additionally, considering fairness among asymmetric users, the researchers also propose an asymptotically fair scheduling algorithm. The technique, opportunistic beamforming and nulling, was introduced in [14]. It is used to induce fast and large fluctuations in the SINR of the users so that multiuser diversity can still be exploited. This technique amplifies the possible multiuser diversity gain while satisfying reasonable latency requirements. In [15], low complexity scheme are considered and compared in means of complexity and performance. It is shown that under quiet simple settings linear beamforming achieves the best tradeoff between performance and complexity. In [16], the authors suggested opportunistic beamforming and scheduling schemes with different linear combining techniques. The authors used EVT to derive throughput and scaling laws of the effective signal-to-interference ratios.

In [17], a scheduling scheme in TDMA system is suggested which uses the Wishart channel matrix HH^\dagger eigenvalues analysis. Another work regarding user selection is [18]. The paper considers the benefit of opportunistic channel state dependent scheduling which can provide significant performance gains for wireless networks by exploiting the independence of fading statistics across the user population. A combination of spatial multiplexing with multiuser diversity is given in [19]. The researchers develop an optimal cross-layer scheduling mechanism that executes fair scheduling at the upper layer and optimal antenna assignment at the physical layer. For fair scheduling, they propose a framework that achieves the objective of maximum capacity and proportional fairness. For optimal antenna assignment, they consider the Hungarian algorithm that maximally utilizes the characteristics of MIMO systems by adopting the graph theoretical approach. More relevant works are [20, 21], and a survey on limits of MIMO BC can be found in [22].

The characteristics of Quality of Service (QoS) parameters can also be used as appropriate scheduling decisions rules, in [23], a multi-user downlink data scheduler with quality-of-service (QoS) provisioning was introduced. In addition to algorithms that are aware of both channel and queue state information to achieve the maximum aggregate utility in the network. Another scheduler using queue-length-based scheduling is introduced in [24], which implements a congestion control mechanism either at the base station or at the end users and can lead to fair resource allocation and queue-length stability. In [25] a practical low-complexity scheduling strategy was suggested by users' queue state, along with sub-optimal multiplexing strategy.

Another relevant work which use order statistics is [26], where the authors suggested a subcarrier assignment algorithm for Orthogonal Frequency Division Multiplexing (OFDM) based systems, and used order statistics to derive an expression for the resulting link outage probability. Order statistics (or EVT) is also required when one wishes to determine the distribution of the selected users, rather than the a priori distribution.

As we can see appropriate user selection is essential for achieving best utilization of the channel capacity.

2.1.2 Multiple-Access Channel

Centralized scheduling systems can use information given to them from the users in order to preform better scheduling decision and pass them along to the users. In [27], a framework for opportunistic scheduling over multiple wireless channels is introduced. Users are selected for data transmission at any time, though with different throughput and system resource requirements. The article introduced and analyzed Multi-channel Fair Scheduler (MFS), the first wireless scheduling algorithm that provide long term deterministic (MFS-D) and probabilistic (MFS-P) fairness guarantees respectively over multiple wireless channels. In [28], the writers present a new low complexity user selection algorithm based on the adaptive Markov chain Monte Carlo (AMCMC) optimization technique to maximize the sum capacity in an uplink MIMO multiuser system. With the proposed algorithm, 99% of the optimal capacity obtained by exhaustive search method can be achieve.

In [29], a novel scheduling metric was derived based on ergodic sum capacity. A per-user receiver-side correlated channel model was combined with a second-order Taylor series of the ergodic sum capacity as a function of SNR. Thus, an ergodic capacity approximation can be written as an explicit function of the channel model parameters, i.e., the receiver-side channel correlation matrices. This approximation can directly serve as a measure of spatial compatibility among a given set of users, i.e., as a scheduling metric. The scheduling metric presented in this work differs from previously proposed metrics mainly in the fact that no instantaneous channel state information is required. Long-term channel statistics are sufficient, implying scheduling robustness and reduced effort in channel sounding. In [30], an effective MIMO precoding scheme to handle the inter-cell interference for uplink communication in a multicell environment is introduced. The precoding scheme allows to select the number of transmit streams adaptively to the surrounding environments.

Distributed Scheduling

In contrast to the centralized approach, distributed scheduling system decisions is preformed autonomously relying solely on information available at the user. The basic assumption is the the users has all information they need in order to preserve the integrity of the algorithm and updates from the BS are negligible.

In [31, 32], a decentralized MAC protocol for Orthogonal Frequency Division Multiple Access (OFDMA) channels is given, it is a threshold based algorithm and each user estimates his channel gain and compares it to the threshold, in case of exceedance the user can transmit. The optimal threshold indeed maximizes the number of slots in which only one user exceeds it. This distributed scheme achieves $1/e$ of the capacity which can be achieved by scheduling the strongest user. [33] extended the distributed threshold scheme to multi-channel setup, where each user competes on m channels and examined proportional fairness in heterogeneous systems. In [34], the authors used a similar approach for power allocation in the multi-channel setup, and suggested an algorithm that asymptotically achieves the optimal water filling solution. To reduce channel contention, [31] [35] introduced a splitting algorithm which resolves collision by allocating several mini-slots devoted to finding the best user. Assuming all users are equipped with a Collision Detection (CD) mechanism, the authors also analyzed the suggested protocol for users not fully backlogged, where the packets randomly arrive with a total arrival rate λ and for channels with memory. [36] used a similar splitting approach to exploit idle channels in a multichannel setup, and showed improvement of 63% compared to the original scheme in [31]. In [37], the authors perform a joint optimization of both

the thresholds and the access probabilities, which provides a fair allocation that achieves a good tradeoff between total throughput and fairness. [38] extended the suggested scheme in [31] and gave a novel way to analyse such threshold-based algorithms, the Point Process Approximation (PPA), and apply it to non identically distributed scenarios as well.

Extreme Value Theorem - one of the ways to analyse such threshold exceedances is to obtain the distribution of the maximum channel gain. A convenient tool for analyzing this kind of extreme distributions is the Extreme Value Theorem (EVT). As the number of users rises ($K \rightarrow \infty$) the distribution function of the maximal capacity degenerates, therefore EVT helps to obtain the limit distribution for large number of *i.i.d.* users. Prior work which is using this tool and refer to these maximal distributions can be found in the article of *Choi and Andrews (2008)* [39], where we can find the scaling laws of maximal base station scheduling via Extreme Value Theory (EVT). By scheduling the user with the strongest channel among K users one can gain a factor of $O(\sqrt{2 \log K})$ in the expected capacity compared to arbitrary (random, Round-Robin) scheduling. Works mentioned in this chapter such as [16, 26, 38], also used EVT for statistical deductions. EVT has been proved useful for analyzing reliable broadcasting as well [40, 41]. In this work EVT is used in the context of stationary process since this work, in contrary to the works presented above, assume the realistic assumption of time dependency.

2.2 Time Dependent Channel

Since this study extends the problem to the case which the users experience a time varying channel, the system model should be described as a system with memory which has different channel states. In the article of *Gillbert (1960)* [42], we can find a model for burst-noise binary channel, where the channel uses a Markov chain in order to illustrate a case of error free with a good state and a case of probability for error in transmission with a bad state. The channel capacity for this model is derived. Proceeding to the article of *Gillbert*, in the article of *elliott (1963)* [43], a means for estimating error rates for binary block codes was introduced. In [44], the researchers form a finite-state Markov channel model to represent Rayleigh fading channels using the expended model of *Gillbert-elliott*. They develop and analyze a methodology to partition the received signal-to-noise ratio (SNR) into a finite number of states according to the time duration of each state. Each state corresponds to different channel quality indicated by bit-error rate (BER). The number of states and SNR partitions are determined by the fading speed of the channel. In [45] a review on the development of fading channel (i.e. channels with states) concerning statistical models, capacity in terms of single user and multiuser transmission and more is presented. More works which are highly relevant are [46, 47] dealing with the capacity of a time dependent channel and the optimal power allocation under average power constraints.

2.3 Limit laws in time dependent channels

The study on the limit laws for *i.i.d.* random variables was fully characterized in many works starting with [48], therefore studies regarding stationary and non stationary sequences and their extreme laws became the natural extension for this problems. One interesting case is the chain dependent process which aligns with our desire to model the users' channel as described above. Chain dependent process was first introduced by Janssen (1969, [49]) and it is known as the $J - X$ process, where we have the bivariate random variables sequence $(J_n, X_n), n \geq 0$. The sequence X_n depends on the state which the process J_n exist in. Exact formulation for the problem will be given later. We will review some relevant work in this subject. In [50] we can find characterization for the limiting behavior of the maximal value distribution of the sequence X_n while conditioning on J_n . With suitable normalization constants, the exact solution to the distribution of the maximum value is given. This solution is computationally difficult and was formed by using the theory of semi-Markov processes, which employs similar formulation for matrices as formed by the $J - X$ process definition, in order to approximate the limiting distribution laws. In addition they establish

basic properties of the normalizing constants, but did not give a specific way to derive them. In [51] and in its preceding work [52] necessary and sufficient conditions are found for the distribution of the maximal value to be relatively and almost surely stable. One important work is [53] which stated that the limiting behaviour of the maximal value is not dependent on the initial distribution of the underlying Markov chain, thus the sequence J_n may start with the stationary distribution of the chain which make X_n also stationary process. In addition the strong mixing property of chain dependent sequences was also given leading eventually to the fact that under fairly general conditions the limiting behavior of the maxima for chain dependent processes is identical to that for the *i.i.d.* process with the same marginal distribution. In [54] and in its preceding paper [55] dependency conditions of the X_n sequence known as D and D' which were given in [56] were used while achieving the results for the limiting distributions for chain dependent process. Properties for a suitable normalizing constants were also given under general conditions, in order for the limiting behaviour will converge to one of the limiting distributions introduced in [48]. Rate of convergence for these approximations for the limiting convergence can be found in [57]. It is important to mention that all the studies above showed convergence to one of the three limiting distribution given in [48] and in some we have specific characterizations for achieving the normalizing constants. In this work the channel model is based on chain dependent process and results regarding the maximal value behavior is relevant for understanding the channel capacity in our scenario, which make the studies above extremely relevant. Similarly to [55] we used the long range and local dependence condition D and D' respectively, for finding the extreme laws for our sequence. However the assumptions made in [55] regarding the marginal distributions are more strict comparing to the ones we did resulting with different normalizing constant for the convergence.

2.4 Queueing system

The analysis of the performance evaluation of our random multiple-access system is an extremely interesting and important theoretical problem. The interaction between the queues in such systems makes the analysis to be inherently difficult [58], and the importance of such analysis was recognized in [59]. As will be described in the sequel, we are dealing with a slotted time system and therefore the most notable system is the slotted aloha. This system was widely studied in the past with two major aspects, the stability region of the system and its QoS properties and performance. Where the first aspect relate to the problem of finding the values of the arrival rates so the system will be stable, and the second aspect relate to the analysis of such system in terms of delay, mean queue size and more. Here we will not engage the stability problem, since existing results are known for the cases of $K = 2$ [60,61] and $K = 3$ [62], and also sufficient bounds on the arrival rates can be found in [60,63,64] for $K > 3$ in order to maintain stability. In fact, we will investigate the performance of such system which have appeared in many works.

Even today there is no clear understanding of the performance of such systems with large number of users (more than 3), due to the complexity in describing the strong interdependence between the user's queues. Only approximate models were developed in order to describe the system behaviour. In [65], an iterative approximation model was suggested while using decoupled Markov chains. A refined model was presented in [3]. In [66], the mean delay was given for the case of two identical users, as well as an approximation for a larger population. Extension for slotted CSMA/CD model can be found in [67].

Chapter 3

Preliminaries

This chapter offers the required technical background which will be used throughout this work. Starting with the Gaussian channel through the definition of a MIMO system and the its capacity under Gaussian channels, statistical tools for analysing independent and dependent sequences of random variables and the review of results regarding sequences which depends on an underlying Markov chain.

3.1 The Gaussian Channel

In information theory the communication between two points means a physical process which the information transfers through it, therefore it is subject to uncontrollable ambient noise and imperfections of the physical signalling process itself. In [68] it is proven that the channel capacity is the tightest upper bound on the amount of information that can be reliably transmitted over a communication channel, while taking into account the interferences.

Our communication system is MIMO, and between each pair of antennas the channel model is a Gaussian channel which is defined as follows:

$$Y = X + Z \quad (3.1.1)$$

Where Y is the output of the channel with input X and with an independent noise Z with variance N . Without any conditions the capacity of this channel may be infinite Since X can take any value. If the variance of the noise is zero, then the receiver receives the transmitted symbol perfectly, whereas if the variance of the noise is non-zero we can choose an infinite subset of inputs so that they are distinguishable at the output with arbitrarily small probability of error. Such a scheme has an infinity capacity as well.

The most common limitation on the input is an energy or power constraint:

$$E[X^2] \leq P \quad (3.1.2)$$

And so the capacity of such channel is:

$$C = \max_{E[X^2] \leq P} \mathcal{I}(X; Y) = \frac{1}{2} \log \left(1 + \frac{P}{N} \right) \quad (3.1.3)$$

The maximum is attained when $X \sim \mathcal{N}(0, P)$.

3.2 Capacity of MIMO

Consider a MIMO system with t transmit antennas and r receive antennas as illustrated in Figure 3.1. The linear link model between the transmit and receive antennas can be represented in the vector notation as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (3.2.1)$$

where \mathbf{y} is the $r \times 1$ received vector from the $t \times 1$ input vector \mathbf{x} . \mathbf{n} is $r \times 1$ zero mean complex Gaussian noise vector with independent, equal variance real and imaginary parts. And \mathbf{H} is a $r \times t$ normalized channel matrix,

$$\mathbf{H} = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,t} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{r,1} & h_{r,2} & \cdots & h_{r,t} \end{pmatrix} \quad (3.2.2)$$

Each element h_{ij} represents the complex gains between the i^{th} transmit and j^{th} receive antennas.

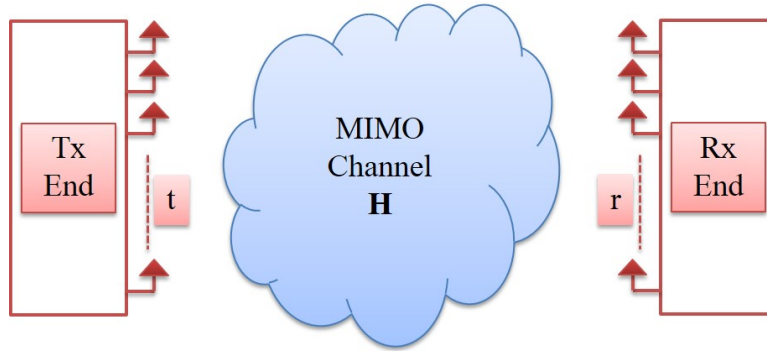


Figure 3.1: MIMO channel

3.2.1 multi-antenna Gaussian channels

After describing the system and the capacity of a single Gaussian channel we would like to explore the capacity of the system. In [4] the researcher investigates the use of multiple antennas at the receiver and at the transmitter for single user communication over the additive Gaussian channel, with or without fading, and calculates the capacity of such a system. The linear model of the system is as given in (3.2.1), where now $\mathbf{y} \in \mathbb{C}^r$, $\mathbf{x} \in \mathbb{C}^t$. The transmitter is constrained in its total power to P ,

$$E[\mathbf{x}^\dagger \mathbf{x}] \leq P. \quad (3.2.3)$$

The matrix \mathbf{H} is a random matrix when all its entries are *i.i.d.* Gaussian variables with zero-mean independent real and imaginary parts, each with variance $1/2$. This matrix models the Rayleigh fading environment of the channel. In [4], the researcher focus in the scenario which the matrix \mathbf{H} , is a random matrix, but in order to expose the techniques used to calculate the channel capacity he consider the case which \mathbf{H} is deterministic.

Using the singular value decomposition and consider \mathbf{H} is deterministic we will derive the capacity of this channel.

This is a vector Gaussian channel, We can compute the capacity by decomposing the vector channel into a set of parallel, independent scalar Gaussian sub-channels using SVD.

The matrix \mathbf{H} can be written as $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger$ where $\mathbf{U} \in \mathbb{C}^{r \times r}$ and $\mathbf{V} \in \mathbb{C}^{t \times t}$ are unitary, and $\mathbf{D} \in \mathbb{R}^{r \times t}$ is non-negative and diagonal (the diagonal elements are the ordered singular values of

the matrix H).

Therefore (4) can be written like this:

$$\mathbf{y} = \mathbf{U} \mathbf{D} \mathbf{V}^\dagger \mathbf{x} + \mathbf{n} \quad (3.2.4)$$

because \mathbf{U} and \mathbf{V} are invertible we can write:

$$\mathbf{U}^\dagger \mathbf{y} = \mathbf{D} \mathbf{V}^\dagger \mathbf{x} + \mathbf{U}^\dagger \mathbf{n} \quad (3.2.5)$$

and we mark it as: $\tilde{\mathbf{y}} = \mathbf{D} \tilde{\mathbf{x}} + \tilde{\mathbf{n}}$, where $\tilde{\mathbf{n}}$ has the same distribution as \mathbf{n} .

Since H is of rank at most $\min\{r, t\}$, at most $\min\{r, t\}$ of the singular values of it are non zero. In addition $\|\tilde{\mathbf{x}}\|^2 = \|\mathbf{x}\|^2$, thus the energy is preserved and we have equivalent representation as a parallel Gaussian channel:

$$\tilde{y}_i = \lambda_i \tilde{x}_i + \tilde{n}_i, \quad 1 \leq i \leq \min\{r, t\} \quad (3.2.6)$$

So the capacity can be calculated as a system of parallel independent channels:

$$C = \sum_i \log \left(1 + \frac{P_i^* \lambda_i^2}{N_0} \right) \quad (3.2.7)$$

where $P_1^*, \dots, P_{\min\{r, t\}}^*$ are the waterfilling power allocations: $P_i^* = (\mu - \frac{N_0}{\lambda_i})^+$, with μ chosen to satisfy the total power constraint $\sum_i P_i^* = P$.

In the case which H is random, it is shown that the capacity of the channel is achieved when \mathbf{x} is a circularly symmetric complex Gaussian with zero-mean and covariance $\frac{P}{t} \mathbf{I}_t$. The capacity is given by

$$E[\log(\det(\mathbf{I}_r + \frac{P}{t} \mathbf{H} \mathbf{H}^\dagger))], \quad (3.2.8)$$

as t grows the expression $\frac{1}{t} \mathbf{H} \mathbf{H}^\dagger \rightarrow \mathbf{I}_r$ almost surely (law of large numbers) and therefore the capacity is

$$r \log(1 + P). \quad (3.2.9)$$

The way to derive the capacity is by maximizing the average mutual information $\mathcal{I}(\mathbf{x}; \mathbf{y})$.

In [5] we can find an overview of the extensive recent results on the capacity of MIMO channels for both single-user and multiuser systems. The great capacity gains predicted for such systems can be realized in some cases, but realistic assumptions about channel knowledge and the underlying channel model can significantly mitigate these gains.

3.2.2 A Gaussian Approximation to the MIMO Capacity

Asymptotic capacity analysis of MIMO channels with arbitrary large number of antennas is widely used to overcome mathematical complexity and often allows for significant insight. Several works [69–71] investigate the accuracy of a Gaussian approximation by deriving the variance and mean of the capacity of a MIMO system resulting with the fact that, distribution of the single-user-MIMO capacity follows the Gaussian distribution with mean that grows linearly with $\min(r, t)$, and variance which is mainly influenced by the power constraint P .

According to this, in this work each user capacity is approximated by a Gaussian distribution, i.e., for each user i , $C_i \sim N(\mu_i, \sigma_i)$.

3.3 Extreme Value Theory

As the number of users raises ($K \rightarrow \infty$) the distribution function of the maximal capacity is degenerating, therefore a need to obtain the limit distribution for asymptotic analysis is required.

Extreme Value Theory (EVT) is a common tool used to approximate such limit distributions, for large number of random variables. In this sub-section, we briefly review some EVT important results that are used in the asymptotic analysis of the capacity gain ([72], [73]).

We first review results for *i.i.d.* sequences. We then elaborate on extremes results for stationary sequences which will be used when analysing the time dependent capacity sequence each user experience.

Theorem 1. ([72], [73]). *Let (X_1, \dots, X_n) be a sequence of i.i.d. random variables with distribution $F(x)$, and let $M_n = \max(X_1, \dots, X_n)$. If there exists a sequence of normalizing constants $a_n > 0$ and b_n such that as $n \rightarrow \infty$,*

$$P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x) \quad (3.3.1)$$

for some non-degenerate distribution G , then G belongs to the type of one of the following three distributions:

$$\begin{aligned} \text{Type I: } G(x) &= e^{-e^{-x}} & -\infty < x < \infty \\ \text{Type II: } G(x) &= \begin{cases} 0 & x \leq 0 \\ e^{-x^{-\alpha}} & \text{for some } \alpha > 0 \quad x > 0 \end{cases} \\ \text{Type III: } G(x) &= \begin{cases} e^{-(-x)^\alpha} & \text{for some } \alpha > 0 \quad x \leq 0 \\ 1 & x > 0 \end{cases} \end{aligned}$$

The three types of distribution limits can be formulated to a generalized extreme value (GEV) distribution type

$$G(z) = \exp\{-(1 + \xi(\frac{z - \mu}{\sigma}))^{-1/\xi}\} \quad (3.3.2)$$

and we say that $F(x)$ is in the domain of attraction of G , where ξ, μ, σ are the shape, location and scale parameters respectively.

In order to determine the shape parameter ξ by the ancestor distribution $F(x)$, let us look on $h(x)$ the following reciprocal hazard function

$$h(x) = \frac{1 - F(x)}{f(x)} \text{ for } x_F \leq x \leq x^F,$$

where $x_F := \inf\{x : F(x) > 0\}$ and $x^F := \sup\{x : F(x) < 1\}$ are the lower and upper endpoints of the ancestor distribution, respectively. Then the shape parameter ξ is obtained as the following limit,

$$\frac{d}{dx} h(x) \xrightarrow{(x \rightarrow x^F)} \xi.$$

We have considering convergence of probabilities of the form $P\{a_n(M_n - b_n) \leq x\}$, which may be rewritten as $P\{M_n \leq u_n\}$ where $u_n = u_n(x) = x/a_n + b_n$. So the theorem from above can be formulated in a different way:

Theorem 2. ([72, Theorem 1.5.1]) *Let $\{X_n\}$ be an i.i.d. sequence. Let $0 \leq \tau \leq \infty$ and suppose that u_n is a sequence of real numbers such that*

$$n(1 - F(u_n)) \rightarrow \tau \text{ as } n \rightarrow \infty.$$

Then

$$P\{M_n \leq u_n\} \rightarrow e^{-\tau} \text{ as } n \rightarrow \infty.$$

For the more interesting and relevant case of normal distribution, we have the following

Theorem 3. ([72, Theorem 1.5.3]) If $\{X_n\}$ are an i.i.d. standard normal sequence of random variables, then the asymptotic distribution of $M_n = \max(X_1, \dots, X_n)$ is a Gumbel distribution. Specifically

$$P\{a_n(M_n - b_n) \leq x\} \rightarrow \exp\{-e^{-x}\} \quad (3.3.3)$$

where

$$a_n = (2 \log n)^{1/2} \quad (3.3.4)$$

and

$$b_n = (2 \log n)^{1/2} - \frac{1}{2}(2 \log n)^{-1/2}[\log \log n + \log 4\pi]. \quad (3.3.5)$$

If $\{X_n\}$ follows the Gaussian distribution with mean μ and variance σ^2 , then the above normalizing constants are

$$a_n = \sigma(2 \log n)^{1/2} \quad (3.3.6)$$

and

$$b_n = \sigma \left[(2 \log n)^{1/2} - \frac{1}{2}(2 \log n)^{-1/2}[\log \log n + \log 4\pi] \right] + \mu. \quad (3.3.7)$$

Please note that if $F(x)$ follows a Gaussian distribution we can see that $\tau = e^{-x}$.

We turn now for results regarding maxima of stationary sequences. Dependence in stationary series can take many different forms, and it is impossible to develop a general characterization of the behavior of extremes unless some constraints are imposed ([73]). The following conditions make precise the notion of extreme events being near-independent if they are sufficiently distant in time.

Definition 1. We say $\{X_n\}$ is strongly mixing if there is a function g on the positive integers with $g(k) \rightarrow 0$ as $k \rightarrow \infty$ such that, if $A \in \mathfrak{F}(X_1, \dots, X_m)$ and $B \in \mathfrak{F}(X_{m+k}, X_{m+k+1}, \dots)$ for some $k, m \geq 1$, then $|P(A \cap B) - P(A)P(B)| \leq g(k)$. Where $\mathfrak{F}(\cdot)$ denotes the σ -field generated by the indicated random variables. Strong mixing was first introduced by Rosenblatt (1956).

In weakening the strong mixing condition, one notes that the events of interest in extreme value theory are typically those of the form $\{X_n \leq u\}$. Hence one may be led to propose a condition like mixing but only required to hold for events of this type.

Definition 2. ([73], [56]) A stationary series X_1, X_2, \dots is said to satisfy the $D(u_n)$ condition if, for all $i_1 < \dots < i_p < j_1 < \dots < j_q$ with $j_1 - i_p > l$,

$$\begin{aligned} &|P_r\{X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n, X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n\} - \\ &P_r\{X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n\}P_r\{X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n\}| \leq \alpha(n, l), \end{aligned}$$

where $\alpha(n, l) \rightarrow 0$ for some sequence l_n such that $l_n/n \rightarrow 0$ as $n \rightarrow \infty$.

If u_n is a sequence that increases with n the long range dependence condition $D(u_n)$ ensures that, for sets of variables that are far enough apart, the difference of probabilities expressed above, while not zero, is sufficiently close to zero to have no effect on the limit laws for extremes.

Another condition that is very relevant to us is the local dependence condition $D'(u_n)$ [74]:

Definition 3.

$$D'(u_n) : \limsup_{n \rightarrow \infty} n \sum_{j=2}^{\lfloor n/k \rfloor} P\{X_1 > u_n, X_j > u_n\} \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

This condition bounds the probability of more than one exceedance among $X_1, \dots, X_{\lfloor n/k \rfloor}$. This will eventually ensure that there are no multiple points in the points process of exceedances which is necessary in obtaining a simple poisson limit for this point process as will be discussed later. A result regarding to the two conditions $D(u_n), D'(u_n)$ is:

Theorem 4. ([72, Theorem 3.4.1]) Let u_n be constants such that $D(u_n), D'(u_n)$ hold for the stationary sequence $\{X_n\}$. Let $0 \leq \tau \leq \infty$. Then

$$P\{M_n \leq u_n\} \rightarrow e^{-\tau} \quad \text{if and only if} \quad n(1 - F(u_n)) \rightarrow \tau.$$

Thus, from the above we can conclude that the particular GEV type which applies is the same as if the sequence $\{X_n\}$ were *i.i.d.* with the same marginal distribution function F , and the same normalizing constants may be used.

3.4 Point Process Approximation

As mentioned earlier the expected channel capacity was derived in a number of method, one of them was done while using the point process approximation. PPA is a tool for analysing threshold arrival rates and tail distributions. This analysis imposes threshold-based algorithm which discussed earlier as an option to achieve the maximum channel capacity. Again we will start with the basic theory which concern an *i.i.d.* sequences and elaborate on stationary process.

A point process on a set \mathcal{A} is a stochastic rule for the occurrence and position of point events ([73]). Our characterization for the statistical properties of the point process is to define a set of non-negative integer-valued random variables, $N(A)$, for each $A \subset \mathcal{A}$, such that $N(A)$ is the number of points in the set A which can represent a period of time $[t_1, t_2]$. The expected number of points in any subset $A \subset \mathcal{A}$, is defined to be the intensity measure $\Lambda(A)$ of the process. The number of points in a given interval follows a Poisson distribution with intensity measure $\Lambda(A)$ and intensity density $\lambda(t, x)$ where:

$$\Lambda(A) = \int_A \lambda(t, x) dA$$

We assume X_1, X_2, \dots, X_n is a sequence of *i.i.d.* random variables, with common distribution function $F(x)$. We suppose that $F(x)$ is in the domain of attraction of some GEV distribution $G(\cdot)$, with normalizing constants a_n and b_n . We then define a sequence of point processes N_1, N_2, \dots on $[0, 1] \times \mathbb{R}^2$ by:

$$N_n = \left\{ \frac{i}{n}, \frac{x_i - b_n}{a_n} : i = 1, \dots, n \right\} \quad (3.4.1)$$

We can notice that the numbers of occurrences counted in disjoint intervals are independent from each other, low value points are normalised to the same value b_l with:

$$b_l = \lim_{n \rightarrow \infty} \frac{x_F - b_n}{a_n}$$

and high value points are retained in the limit process. Under all the above we get:

Theorem 5. ([75], [73]) The sequence N_n on the set $[0, 1] \times (b_l + \epsilon, \infty)$, where $\epsilon > 0$, converge to a non-homogeneous Poisson process N , that is:

$$N_n \rightarrow N \quad \text{as} \quad n \rightarrow \infty,$$

with intensity:

$$\lambda(t, x) = (1 + \xi x)_+^{-1-1/\xi}$$

We are interesting in the sets of the form: $B_u = [0, 1] \times (u, \infty)$ where $u > b_l$ which reflect the

users with a capacity that exceeded a threshold u . In this case the intensity measure is:

$$\begin{aligned}
\Lambda(B_u) &= \Lambda([0, 1] \times (u, \infty)) \\
&= \int_{t=0}^1 \int_{x=u}^{\infty} \lambda(t, x) dx dt \\
&= \int_{t=0}^1 \left[- (1 + \xi x)_+^{-1/\xi} \right]_{x=u}^{\infty} dt \\
&= \int_{t=0}^1 (1 + \xi u)_+^{-1/\xi} dt \\
&= (1 + \xi u)_+^{-1/\xi}
\end{aligned}$$

where a_+ denotes $\max\{0, a\}$. The occurrences of capacity exceedances of the threshold u can be modeled by a Poisson process, with parameter $\Lambda(B_u)$. In case the sequence $\{X_n\}$ are standard normal random variables:

$$\lim_{\xi \rightarrow 0} (1 + \xi u)_+^{-1/\xi} = e^{-u} = \tau \Big|_{x=u}$$

Since we are interesting in time dependent environment we would like to explore the point process analysis for stationary sequences. In [76] we can find the point process of exceedances of a high level u_n by a stationary sequence $\{X_i\}$ (i.e., points where $X_i > u_n$), obtaining Poisson limits under quite weak dependence restrictions. These involve a long range dependence condition $D(u_n)$ and a local dependence condition $D'(u_n)$ as defined in definitions 2 and 3.

Theorem 6. ([76]) *Let $D(u_n), D'(u_n)$ hold (u_n satisfying $u_n = u_n(\tau)$, such that $1 - F_1(u_n) = P\{X_1 > u_n\} \sim \tau/n$ as $n \rightarrow \infty$ for all $\tau > 0$) for the stationary sequence $\{X_i\}$. Let N_n be the point process, consisting of the exceedances of $u_n(\tau)$. Then $N_n \xrightarrow{d} N$ as $n \rightarrow \infty$, where N is a Poisson process with parameter τ .*

As a result of the above theorem, the point process under the dependence consideration behaves just like one obtained from an *i.i.d.* sequence with the same marginal d.f. F . If the local condition is weakened or omitted, then clustering of exceedances may occur.

3.5 Chain Dependent Process

In this work we use a specific model for dependent data referred to as chain dependent model. We will review the formulation for this model and some relevant results.

Let $\{J_n\}$ be a finite state space Markov chain with ergodic probability transition matrix $P = (p_{ij})$, $1 \leq i, j \leq m$. Further let $\{X_n\}$ be a chain dependent sequence such that

$$\begin{aligned}
P(J_n = j, X_n \leq x \mid X_0, J_0, \dots, X_{n-1}, J_{n-1} = i) \\
= P(J_n = j, X_n \leq x \mid J_{n-1} = i) = p_{ij} H_i(x)
\end{aligned} \tag{3.5.1}$$

where H_i , $i = 1, \dots, m$ are the underlying non-degenerate and honest ($H_i(+\infty) = 1$) distributions, associated with each state of the chain.

Immediate consequences [50]:

- i The stationary probabilities associated with P are (π_1, \dots, π_m) .
- ii $P(X_n \leq x \mid J_{n-1} = i) = H_i(x)$.
- iii $P(X_1 \leq x_1, \dots, X_n \leq x_n \mid J_0, J_1, \dots, J_{n-1} = i) = \prod_{i=1}^n P(X_i \leq x_i \mid J_{i-1})$
The random variables $\{X_n\}$ are conditionally independent given the chain in precisely the sense given by (iii).

Define a nonnegative $m \times m$ matrix by $Q(x) = (p_{ij}H_i(x))$ and the random variable $M_n = \max(X_1, \dots, X_n)$, then In [50] we can find the distribution of the maximum given by

$$P(J_n = j, M_n \leq x \mid J_0 = i) = Q_{ij}^n(x) \quad (3.5.2)$$

where Q^n denotes n th matrix power of Q . The matrix Q with suitable normalizing constants $a_{ijn} > 0$ and b_{ijn} , $i, j = 1, \dots, m$, $n \geq 1$, convergence to non-degenerate mass matrix as $n \rightarrow \infty$

$$\{Q_{ij}^n(a_{ijn}x + b_{ijn})\} \rightarrow \{U_{ij}(x)\} \quad \text{as } n \rightarrow \infty.$$

The Perron-Frobenius eigenvalue $\rho_U(x)$, of $\{U_{ij}(x)\}$, is an extreme value distribution. To attain this result the Semi-Markov matrices theory was used.

A very important results given in [53] states that that the limiting behaviour of M_n is not dependent on the initial distribution of the underlying Markov chain. Thus it may be started with the stationary distribution ($J_0 = \pi_i$, $i = 1, \dots, m$), in which case $\{X_n\}$ is stationary. In addition it is stated that every stationary chain-dependent process $\{X_n\}$ is strongly mixing (Lemma 1.1, [53]). Then under any initial distribution for J_0 (theorem 3.1, [53]) we have

$$P(M_n \leq d_n) \rightarrow l \quad \Leftrightarrow \quad H^n(d_n) \rightarrow l \quad \text{as } n \rightarrow \infty \quad (3.5.3)$$

where d_n is a sequence of reals, $0 < l < 1$ and $H(x)$ is a distribution function such that $H(x) = \sum_{i=1}^m \pi_i H_i(x)$.

The dependency of the sequence $\{X_n\}$ can be characterized such that it would be easier to attain the limiting distribution of M_n . Such dependency conditions are the $D(u_n)$ and $D'(u_n)$ as given in definitions 2 and 3. In [55] a similar analysis was performed for this purpose. They defined a class of sequences, where the chain dependent process is a sub case of it, that fulfill some properties which under them the extreme laws could be achieved. Among those properties, the ones that are relevant to us are that the $D(u_n)$ and $D^*(u_n)$ (an extension for $D'(u_n)$) conditions hold and that there exists a sequence of real numbers $\{u_n\}$ and real numbers z_i , $i = 1, \dots, m$, such that for every H_i , $i = 1, \dots, m$, as in (ii) we have

$$1 - H_i(u_n) = \frac{z_i}{n} + o\left(\frac{1}{n}\right) \quad \text{as } n \rightarrow \infty. \quad (3.5.4)$$

Under the conditions (1a) to (1e) in [55] their basic result is

Theorem 7. ([55]) *Let $\{X_n\}$, $\{u_n\}$ be sequences of random variables and real numbers respectively, satisfying the conditions (1a) to (1e) given. Then*

$$\lim_{n \rightarrow \infty} P(M_n \leq u_n) = e\left\{-\sum_{i=1}^m p_i z_i\right\} \quad (3.5.5)$$

where p_i , $i = 1, \dots, m$, are the stationary distribution on the chain and z_i are as given in (3.5.4).

In addition for the above theorem they give a corollary regarding the properties for convergence to the specific extreme value distributions. However assumption (2.2) in their work does not hold for this study marginal distributions, which associated with the states of the chain. Therefore their specific rules for convergence does not apply here exactly.

Chapter 4

Model Description

We consider a slotted MIMO uplink system with K independent users and one base station. The time axis is divided into fixed-length intervals referred to as time slots or, simply, slots, in the sequel. The distribution of the channel capacity a single user experiences, is time varying according to a Gilbert Elliott model [42]. We first assume that a single user, with the strongest received signal at the base station, is selected by a centralized scheduler to utilize the channel and transmit in each slot. We then consider the distributed setting as well. Hence, we first wish to analyze the expected channel capacity where the base station chooses the strongest user for transmission using the CSI sent to it from the users. We then turn to analyzing the capacity under a distributed opportunistic user selection algorithm, which selects the strongest user to transmit, without coordinating between the users and without all users sending CSI to the base station. We assume the transmitters approximate their channel parameters from a pilot signal sent from the receiver.

The linear model of the system is:

$$\mathbf{y}(n) = H(n)\mathbf{x}(n) + \mathbf{N}(n) \quad (4.0.1)$$

$\mathbf{y} \in \mathbb{C}^r$ is the received vector from the input vector $\mathbf{x} \in \mathbb{C}^t$ where H is a $r \times t$ complex matrix (r - number of receiving antennas, t - number of transmitting antennas) and \mathbf{N} is zero mean *i.i.d.* complex Gaussian noise with independent, equal variance real and imaginary parts. The transmitter is constrained in its total power to P ,

$$E[\mathbf{x}^\dagger \mathbf{x}] \leq P.$$

The matrix $H(n)$ is a random matrix which models the Rayleigh fading environment of the time dependent channel.

4.1 Time dependent channel

the channel distribution each user sees is governed by a Markov chain with two states, G (for Good) and B (for bad), where each state determines the channel distribution (Figure 4.1). Thus, the channel is memoryless conditioned on the state process, and the capacity of the i -th user in each slot is determined by the Good Bad Markov process $\{J(n)\}$:

$$C_i(n) = \begin{cases} N_g & \text{when } J_i(n) = \text{Good} \\ N_b & \text{when } J_i(n) = \text{Bad}, \end{cases} \quad (4.1.1)$$

where N_g and N_b are normal random variables with parameters (μ_g, σ_g) and (μ_b, σ_b) , respectively. This is due to the Gaussian approximation for the MIMO channel capacity [70, 71]. The parameters reflect the differences in the channel qualities, that is, good channel parameters maintain

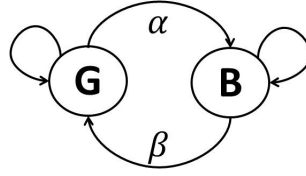


Figure 4.1: A Good Bad channel model according to [42]. The capacity distribution is either good or bad Gaussian random variable according to the state of the system.

(1) $\sigma_g > \sigma_b$, $\mu_g, \mu_b \in \mathbb{R}$ or (2) $\sigma_g = \sigma_b$ and $\mu_g > \mu_b$.

As mentioned, we assume a non trivial arrival process and that the users are not always backlogged. The arrival process of new packages to the system is characterized by a Poisson process. In addition, we assume that all user are homogenous, thus, all users has the same arrival rate λ . We present several approximate models. Thus, new packages may enter the system at any given (continuous) point on the time axis or at the beginning of a slot, depending on the approximate model. The size of the packages is set such that the transmission time is exactly one slot. We define the service time as the time from the moment a package becomes first in queue, until it is successfully transmitted. Hence, the service time of the packages is composed of the waiting time for transmission and the transmission time. In addition, we assume that services are synchronized to (i.e., can only start and end at) slot boundaries. This of course may cause collisions.

Chapter 5

Capacity under time dependent channel

In order to exploit user diversity, the user with the best channel must utilizes the channel. Therefore, the problem is in finding the distribution of the random variable \widetilde{M}_K , defined as:

$$\widetilde{M}_K = \max\{C_1(n), C_2(n), \dots, C_K(n)\}, \quad (5.0.1)$$

where $C_i(n)$ is defined in (4.1.1). We are interested in the expected channel capacity $E[\widetilde{M}_K]$.

In the beginning of each slot the channel each user experience depends on the state he exists in, resulting with two independent groups of users, a good group and a bad group. The two groups are independent due to the independency between the users, and in each separate group the users identically distributed. The size of these groups is constantly changes according to each user state process which happen in the beginning of each slot. Given the state of all users the groups size are known and we may derive the expected channel capacity in the following way

$$\begin{aligned} E[\widetilde{M}_K] &= E[\max\{C_1(n), C_2(n), \dots, C_K(n)\}] \\ &= \sum_{j=0}^K E\left[\max\left\{\max\{C_1(n), \dots, C_j(n)\}, \max\{C_{j+1}(n), \dots, C_K(n)\}\right\} \middle| j\right] P(j) \end{aligned} \quad (5.0.2)$$

The distribution of the maximum for each group can be obtain by using EVT, although the complicated part is to do so for every group size. Unfortunately due to complicated analytical expressions the result remained open, and more details can be found in the appendices regarding this way of analysis.

Nevertheless we were able to achieve a closed form for the channel capacity in more informative ways. The first one is by using the stationarity of the users' capacities sequences, and the second is to analyse the rate of exceedance above some threshold while imposing our distributed algorithm.

5.1 Capacity - Centralized Scheduling

In this section we give the expected channel capacity with the assumption that the strongest user utilizes the channel. Let us consider the distribution of each $C_i(n)$, as defined in (4.1.1), which is determined by the Good-Bad Markov process $\{J(n)\}$. The stationary distribution of the chain is $(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta})$, which we will denote as (p, q) . Therefore, the distribution of $C_i(n)$ assuming a

stationary distribution is as follows:

$$\begin{aligned}
F(t) &= P(C_i(n) \leq t) = \\
&= P(N_g(n) \leq t \mid J(n) = G)P(J(n) = G) + \\
&= P(N_b(n) \leq t \mid J(n) = B)P(J(n) = B) \\
&= pF_g(t) + qF_b(t),
\end{aligned} \tag{5.1.1}$$

where $F_g(t)$ and $F_b(t)$ are Gaussian distributions with parameters (μ_g, σ_g) and (μ_b, σ_b) respectively which represent the users' different channel states. The distributions parameters maintain $\sigma_g > \sigma_b$, $\mu_g, \mu_b \in \mathbb{R}$ or $\sigma_g = \sigma_b$ and $\mu_g > \mu_b$. The distribution of the maximal value is in the form of

$$P(\widetilde{M}_K \leq x) = P(C_1(n) \leq x, C_2(n) \leq x, \dots, C_n(n) \leq x) = F^n(x),$$

due to the independence between the users. We wish to test the behavior of $F^n(x)$ as $n \rightarrow \infty$.

The legitimacy to use the stationary distribution also comes from [53] as given in (3.5.3), the convergence of $F^n(d_n)$ to a real number l , where $0 < l < 1$ and d_n is a sequence of reals, exist if and only if, $P(M_n \leq d_n) \rightarrow l$ as $n \rightarrow \infty$. Thus we can analyze the expected channel capacity using EVT with the distribution above, as the marginal distribution function for each user. In order to do so, first we need to prove that convergence to one of the extreme distributions types $G(x)$ exists, and then find the constants a_n and b_n such that as $n \rightarrow \infty$, $P\{\widetilde{M}_K \leq u_n\} \leq x \rightarrow G(x)$. [72, 1.6.1, 1.6.2] gives necessary and sufficient conditions on the marginal distribution F to belong to each of the three possible domains of attraction of the extreme value distributions. The first sufficient condition states that if f has a negative derivative f' for all x in some interval (x_0, x_F) , $(x_F \leq \infty)$, $f(x) = 0$ for $x \geq x_F$, and

$$\lim_{t \uparrow x_F} \frac{f'(t)(1 - F(t))}{f^2(t)} = -1 \tag{5.1.2}$$

then F is in the domain of attraction of Type I extreme value distribution (Gumbel distribution). The second necessary and sufficient condition for F to be in the domain of attraction of Type I extreme value distribution, states that there exists some strictly positive function $g(t)$ such that

$$\lim_{t \uparrow x_F} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x} \tag{5.1.3}$$

for all real x .

Our main result in this context is the following.

Theorem 8. *Let (C_1, \dots, C_K) be a sequence of random variables with distribution $F(t)$ as given in (5.1.1), then the asymptotic distribution of $\widetilde{M}_K = \max(C_1, \dots, C_K)$ is of Type I. Specifically,*

$$P\{a_K(\widetilde{M}_K - b_K) \leq x\} \rightarrow e^{-e^{-x}} \tag{5.1.4}$$

where

$$a_K = \frac{\sqrt{2 \log K}}{\sigma_g} \tag{5.1.5}$$

and

$$b_K = \sigma_g \left(\sqrt{2 \log K} - \frac{\log \log K + \log \frac{4\pi}{p^2}}{2\sqrt{2 \log K}} \right) + \mu_g. \tag{5.1.6}$$

The proof outline is as follows. First we start by solving directly the limit in condition (5.1.2), while using the sandwich rule, replacing the Gaussian CDF with its corresponding complimentary error function and using the upper and lower bounds of it. The second condition (5.1.3) has been

proven by finding that $g(t) = \sigma^2/(t - \mu_g)$, along with the asymptotic analysis that the "Bad" Gaussian distribution has little affect in it's tail comparing to the "Good" distribution. And finally using similar asymptotic analysis as preformed in condition (5.1.3) the normalizing constants was derived. For more details, see appendix A, which gives the fulfillment of the two conditions above and the derivation of the normalizing constants a_K and b_K .

Using Theorem 8 we can now give the expected channel capacity

$$\begin{aligned} E[\widetilde{M}_K] &= b_K + \frac{\gamma}{a_K} \\ &= \sigma_g \left(\left(\sqrt{2 \log K} - \frac{\log \log K + \log \frac{4\pi}{p^2}}{2\sqrt{2 \log K}} \right) + \frac{\gamma}{\sqrt{2 \log K}} \right) + \mu_g \end{aligned} \quad (5.1.7)$$

where $\gamma = 0.57721$ is Euler-Mascheroni constant. Thus, for K sufficiently large,

$$E[\widetilde{M}_K] = \sigma_g \sqrt{2 \log K} + \mu_g + o\left(\frac{1}{\sqrt{\log K}}\right). \quad (5.1.8)$$

That is, for large number of users, the expectation capacity grows like $\sigma_g \sqrt{2 \log K}$. A word is in place here concerning the values of p and q . We assume that (p, q) are bounded away from zero, which implies that the uninteresting case which all users has bad channel state, or the opposite case where all have good channel does not apply. Also, in the proof of Theorem 8 we used this assumption while we made a division by p and q . Simulation results for the capacity distribution is given in Figure 5.1.

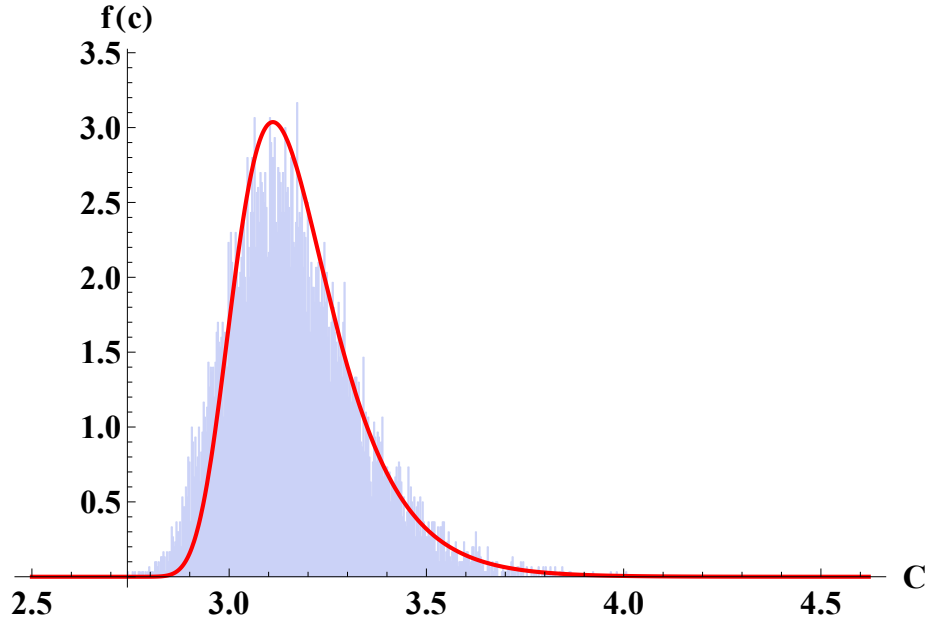


Figure 5.1: Simulation for the maximal capacity distribution, when choosing the maximal capacity among 5000 capacities that following the stationary distribution where the red line is the corresponding Gumbel density with the constants a_K and b_K .

Remark 1 (The probability p). *The probability p , which is the stationary probability to exist in a good state, rules the average number of users which are in good state while considering the whole system. In each time slot we can distinguish between two groups of users, the users that are in good state and the users that are in bad state. It is easy to show that the number of good users in the*

good group on average is pK . Hence, as p grows the expected capacity grows since there are more users in good state, we can see it in the analytical result (5.1.7) for the channel capacity and in Figure 5.2.

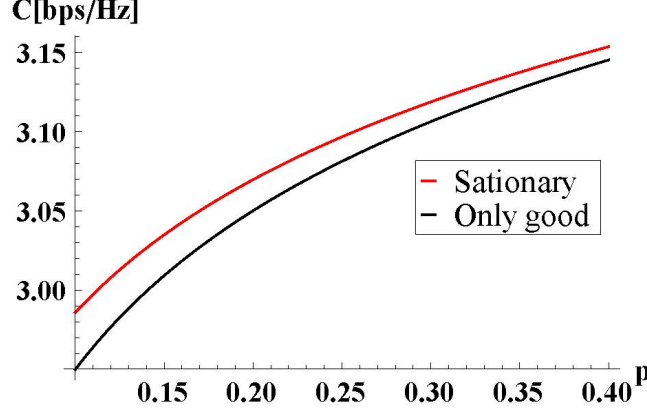


Figure 5.2: The channel capacity gain as given in (5.1.7) in comparison to the gain of choosing only from the good group with size Kp , for 5000 users as a function of p and $\mu_g = \sqrt{2}, \sigma_g = 0.5$.

One may wonder on why considering the bad group at all, meaning, why should a user which belongs to the bad group will be taken into account in the scheduling decision process in the beginning of a slot. Thus leaving only the users with the good channel to compete for the channel. For this case, the capacity with K sufficiently large is

$$E[\widetilde{M_{pK}}] = \sigma_g \sqrt{(2 \log pK)} + \mu_g + o\left(\frac{1}{\sqrt{\log pK}}\right) \quad (5.1.9)$$

In Figure 5.2 we can see the influence of the bad group on the capacity gain. For rather small values of p it is beneficial to schedule the strongest user from both groups. As p grows the size of the good group grows as well hence the two capacity gains converge to the case were all the users are in good state. In Figure 5.3 we can see the capacity gain as a function of the number of users, specifically, in Figure 5.3b the difference between the gains is noticeable.

5.2 Capacity - Distributed Scheduling

When considering a distributed scheduling algorithm, we derive the expected channel capacity using PPA. PPA is a tool for analysing threshold arrival rates and tail distributions, while imposing a threshold-based algorithm for the channel access, which was discussed earlier as an option for a distributed scheduling mechanism. This method can also be found in [77] for the case of *i.i.d.* and heterogeneous users. First we analyze each user separately and examining its sequence of channel capacities over time. This sequence can be modeled as a point process on some interval, using the definition in (3.4.1), yet with the different that our process of points is formed from a dependent sequence of random variables. As it turns out, the exceeding points above some threshold u converges to a non-homogeneous Poisson process, which we discuss later, and the question that can be asked is, what is the rate of exceedance above the threshold u . We show that our dependent sequence converge to the same Poisson process if the sequence was of an *i.i.d.* random variables, which will eventually allow us to examine the sequence of all users' capacities in a specific time.

The sequence $\{C_i(n)\}$ depends on what state user i exists in, which is, in turn, ruled by the underlying Markov chain process $\{J_i(n)\}$ as defined in (4.1.1). This kind of process have been studied before and is known as a $\tilde{J} - X$ process by Janssen (1969, [49]), or “chain-dependent” process,

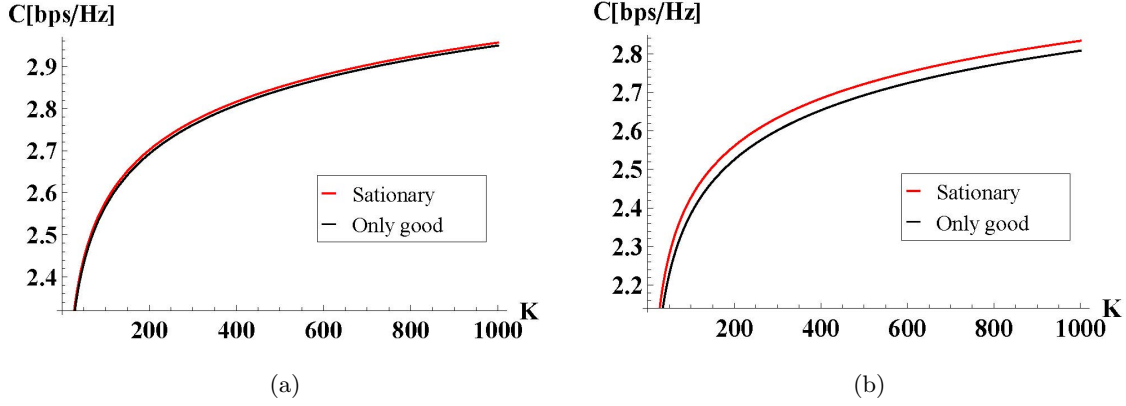


Figure 5.3: The capacity gain for choosing the best user among all the population and only the good where $\sigma_g = 0.5$ and (a) $p=0.5$ (b) $p=0.2$

emphasizing the fact that this is a natural extension of an *i.i.d.* process. The extreme laws for chain dependent processes have showed that, convergence to one of the extreme distributions exist under some conditions. This is important for the formation of the point process and the analysis of the rate of exceedance we do in the sequel.

Before we give the formulation for the chain dependent process in our work, let us remind the formulation of the original $\tilde{J} - X$ process as reviewed also in Chapter 3.5. The $\tilde{J} - X$ process as in [78], [53] satisfies:

$$\begin{aligned} P(\tilde{J}_n = j, X_n \leq \alpha \mid \tilde{J}_0, X_1, \tilde{J}_1, \dots, X_{n-1}, \tilde{J}_{n-1} = i) \\ = P(\tilde{J}_n = j, X_n \leq \alpha \mid \tilde{J}_{n-1} = i) = P_{ij} \tilde{H}_i(\alpha), \end{aligned} \quad (5.2.1)$$

where P is the transition matrix of the chain and $\tilde{H}_i(\alpha)$, where i belongs to the state space, are the distribution functions associated with the chain states, respectively. Note that each state determines the distribution of X for the preceding time transition. This means that given the chain process $\{\tilde{J}_n\}$ The random variables of the $\{X_n\}$ process are conditionally independent. More precisely,

$$\begin{aligned} P(X_1 \leq \alpha_1, X_2 \leq \alpha_2, \dots, X_n \leq \alpha_n \mid \\ \tilde{J}_0 = j_0, \tilde{J}_1 = j_1, \dots, \tilde{J}_{n-1} = j_{n-1}) \\ = P(X_1 \leq \alpha_1 \mid \tilde{J}_0 = j_0) P(X_2 \leq \alpha_2 \mid \tilde{J}_1 = j_1) \cdots \\ P(X_n \leq \alpha_n \mid \tilde{J}_{n-1} = j_{n-1}) \\ = \tilde{H}_{j_0}(\alpha_1) \tilde{H}_{j_1}(\alpha_2) \cdots \tilde{H}_{j_{n-1}}(\alpha_n). \end{aligned}$$

As discussed in Chapter 3.5, every stationary chain-dependent process $\{X_n\}$ is strongly mixing. This is true if the initial distribution of the chain is the stationary distribution, i.e. $P(J_0 = i) = \pi_i$ for all i in the finite state space, where π is the stationary distribution. Then, the distribution of X_n is $\tilde{H}(x) = \sum \pi_i \tilde{H}_i(x)$. This result is very important to us since we are interested in a time dependent environment and we would like to explore the point process analysis for the stationary sequences associated with it.

Returning to our context, we shall consider the sequence $\{X_n\}$ as the sequence of capacities $\{C_i(n)\}$ of the i -th user over time and the chain process $\{\tilde{J}_n\}$ as the sequence of the irreducible, aperiodic, 2-state Good-Bad Markov chain $\{J_i(n)\}$. The distributions $\tilde{H}_i(\alpha)$ will be the distributions $H_i(\alpha)$, where $i = g, b$, according to the Good-Bad states, respectively. Since we are analysing the capacity process of one user, which is identically to the other users, we will note $\{C_n\}$ as it's capacity sequence and $\{J_n\}$ as the sequence of the chain.

Unlike the definition of the $\tilde{J} - X$ process, in our model each state determines the capacity distribution in the current slot. So equation (5.2.1) according to our case can be written:

$$\begin{aligned} P(J_n = j, C_n \leq \alpha \mid J_0, C_1, J_1, \dots, C_{n-1}, J_{n-1} = i) \\ = P(J_n = j, C_n \leq \alpha \mid J_{n-1} = i) = P_{ij} H_j(\alpha) \end{aligned} \quad (5.2.2)$$

The fact that the capacity in each slot depends on the chain state in the current slot doesn't impact the strong mixing property of the sequence $\{X_n\}$. We will show that the $J - C$ process as defined to meet our model, is still strongly mixing by definition 1, while using the proof guidelines in [78].

Lemma 1. $\{C_n\}$ is strongly mixing with $g(k) = \sum_j \pi_i \mid P_{ij}^k - \pi_j \mid$, where π is the stationary distribution of the chain and $P_{ij}^k = (P^k)_{ij}$.

Proof. Let A and B be as definition 1. Then

$$\begin{aligned} & \mid P(A \cap B) - P(A)P(B) \mid \\ & \leq \sum_{i,j \in \{0,1\}} \mid P(A \cap B, J_m = i, J_{m+k} = j) - \\ & \quad P(A, J_m = i)P(B, J_{m+k} = j) \mid \\ & = \sum_{i,j \in \{0,1\}} P(A \mid J_m = i)P(B \mid J_{m+k} = j)P(J_m = i) \\ & \quad \mid P(J_{m+k} = j \mid J_m = i) - P(J_{m+k} = j) \mid \\ & \leq \sum_{i,j \in \{0,1\}} \pi_i \mid P_{ij}^k - \pi_j \mid = g(k). \end{aligned}$$

For each $i \in 0, 1$, $\sum_j \pi_i \mid P_{ij}^k - \pi_j \mid \rightarrow 0$ as $k \rightarrow \infty$, this was also used in [78] for showing the strongly mixing property, nevertheless, it is easy to notice that in fact as $k \rightarrow \infty$, $(P^k)_{ij} \rightarrow \pi_j$ regardless to i . \square

Considering Theorem 6, we will show that the point process under the chain dependent model convergence to a Poisson process just like one obtained from an *i.i.d.* sequence with the same marginal d.f. F , which in our case is $H(x)$. The normalizing constants will be as given in Theorem 8 which maintains the condition $1 - H(u_n) = P\{X > u_n\} \sim \tau/n$ as $n \rightarrow \infty$. As mentioned before, the strong mixing condition hold for our sequence of channel capacity, and as a result so does the condition $D(u_n)$ as defined in definition 2 which is a weakened case of strong mixing [56]. We would like to show that condition $D'(u_n)$ as defined in definition 3 also holds so we would be able to characterize the rate of exceedance over the threshold u_n . Note that we are only interested in a sequence of reals $\{u_n\}$ which satisfies $1 - H(u_n) = \tau/n + o(1/n)$, in the fulfillment of the $D'(u_n)$ condition.

Lemma 2. The local dependence condition $D'(u_n)$ holds for the sequence $\{C_n\}$ as defined in (4.1.1).

Proof.

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} n \sum_{r=2}^{\lfloor n/k \rfloor} P(C_1 > u_n, C_r > u_n) \\
& \stackrel{(a)}{\leq} \limsup_{n \rightarrow \infty} n \sum_{r=2}^{\lfloor n/k \rfloor} \sum_{i,j \in \{0,1\}} P(C_1 > u_n | J_1 = i) \\
& \quad \cdot P(C_r > u_n | J_r = j) P(J_1 = i) P(J_r = j | J_1 = i) \\
& = \limsup_{n \rightarrow \infty} n \sum_{r=2}^{\lfloor n/k \rfloor} \sum_{i,j \in \{0,1\}} (1 - H_i(u_n))(1 - H_j(u_n)) \pi_i P_{ij}^r \\
& \stackrel{(b)}{\leq} \limsup_{n \rightarrow \infty} n \sum_{r=2}^{\lfloor n/k \rfloor} \sum_{i,j \in \{0,1\}} \left(\frac{\tau}{n\pi_i} + o\left(\frac{1}{n}\right) \right) \\
& \quad \cdot \left(\frac{\tau}{n\pi_j} + o\left(\frac{1}{n}\right) \right) \pi_i P_{ij}^r \\
& \leq \limsup_{n \rightarrow \infty} n \sum_{r=2}^{\lfloor n/k \rfloor} \sum_{i,j \in \{0,1\}} \frac{1}{\pi_j \pi_i} \left(\frac{\tau}{n} + o\left(\frac{1}{n}\right) \right)^2 \pi_i P_{ij}^r \\
& \leq \limsup_{n \rightarrow \infty} n \left\lfloor \frac{n}{k} \right\rfloor \left(\frac{\tau}{n} + o\left(\frac{1}{n}\right) \right)^2 \sum_{i,j \in \{0,1\}} \frac{1}{\pi_j \pi_i} \pi_i \max_r \{P_{ij}^r\} \\
& = (\tau^2 + o(1))^2 \frac{1}{k} \sum_{i,j \in \{0,1\}} \frac{1}{\pi_j \pi_i} \pi_i \max_r \{P_{ij}^r\} \rightarrow 0 \text{ as } k \rightarrow \infty
\end{aligned}$$

In the above chain, (a) is since once J_1 is known, C_1 and C_r, J_r are independent, then we conditioned on J_r . (b) is true since u_n maintains $n(1 - F_1(u_n)) = \tau + o(1)$ so we have

$$\begin{aligned}
1 - F_1(u_n) &= 1 - H(u_n) = 1 - \sum_i \pi_i H_i(u_n) \\
&= \sum_i \pi_i (1 - H_i(u_n)) = \frac{\tau}{n} + o\left(\frac{1}{n}\right)
\end{aligned}$$

and therefore

$$\pi_l (1 - H_l(u_n)) = \frac{\tau}{n} - \sum_{i, i \neq l} \pi_i (1 - H_i(u_n)) \leq \frac{\tau}{n} + o\left(\frac{1}{n}\right).$$

Note that π_l and τ are constants. □

Thus, in our paradigm a single users' channel capacity process has the same laws of convergence as if the sequence $\{C_n\}$ was *i.i.d.* with marginal distribution $H(x)$. Therefore, since the users are independent and each user sees the same marginal distribution $H(x)$, we can analyze the point process of the sequence of all users' capacities at a specific time (e.g. time slot), resulting in the basic case of *i.i.d.* random variables, as given in Theorem 5. Considering the above, we now turn to evaluate the expected channel capacity.

5.2.1 Distributed Algorithm

The distributed algorithm for the channel access is based on the algorithm suggested in [32]. Given the number of users, we set a capacity threshold u such that only a small fraction of the users will exceed it. At the beginning of each slot, each user estimates its capacity for that slot. If the capacity anticipated by the user is greater than the capacity threshold, it transmits in that slot.

Otherwise, it keeps silent. Choosing the appropriate threshold value is of great significance. A lower value means more users exceedances, therefore more collisions, while a high value means more idle slots. The optimal threshold value is set such that one user exceeded the threshold on average, and as a result, its transmission is successful. We treat this slot as a utilized slot. Hence the expected channel capacity has the form:

$$C_{av}(u) = P_r(\text{utilized slot})E[C|C > u]$$

In order to calculate $E[C|C > u]$, the expected capacity experienced by a user who passed u , it is important to understand that the average capacity is ruled by the values of points that exceeds the threshold, so one needs to evaluate the distance of the exceeding points from the threshold. [75] give the analytical tools to compute the tail distribution of the exceeding points. These points follow the generalized Pareto distribution, hence by using the PPA exceedance rate results, and since we showed that we have the same exceedance as an *i.i.d.* case with $H(x)$ as the marginal distribution, the result is the same and we have

$$E[C|C > u] = u + \frac{1}{a_K} + o\left(\frac{1}{a_K}\right),$$

where a_K is the normalizing constant as in Theorem 8.

We say that a slot is utilized if only one point out of all K points exceeds the threshold. Hence, as $K \rightarrow \infty$

$$K \left(\frac{1}{K}\right) \left(1 - \frac{1}{K}\right)^{K-1} \rightarrow e^{-1},$$

where the threshold u was chosen such that $1 - H(u) = 1/K$. We will elaborate on this value in the next subsection. The expected channel capacity thus

$$C_{av}(u) = e^{-1} \left(u + \frac{1}{a_K} + o\left(\frac{1}{a_K}\right) \right) \quad (5.2.3)$$

In the next subsection we show that $u = b_K$ so when comparing the capacity result in (5.2.3) to the expected capacity $E[\widetilde{M}_K]$ as given in (5.1.7), for the centralized approach, we see that we have the same scaling laws, only that the distributed approach is smaller only by a factor of e^{-1} . The normalizing constant a_K along with the threshold $u = b_K$ is as given in Theorem 8, reflect the time dependency of the channel as explained before.

5.2.2 Threshold Estimation

The optimal threshold was set such that only one user on average exceeds it in each time slot. This is due to the analysis of [31], where it has been shown that the optimal transmission probability for each user to transmit in a specific time slot is $\alpha(n)/n$, where asymptotically $\alpha(n) \rightarrow 1$ hence, the probability $1/n$ is a good approximation. By using this rule we can estimate the optimal threshold value for the time dependent channel as well:

$$1 - H(u) = \frac{1}{K}$$

Hence,

$$1 - pF_g(u) - qF_b(u) = \frac{1}{K},$$

where the above represent the probability that a user capacity will exceed the threshold u . With similar derivation as in appendix A we get that

$$\begin{aligned} u &= \sigma_g \sqrt{2 \log K} \left(1 - \frac{\frac{1}{2} \log \frac{4\pi}{p^2} + \frac{1}{2} \log \log K}{2 \log K} + o\left(\frac{1}{\log K}\right) \right) \\ &\quad + \mu_g \\ &= b_K \end{aligned} \quad (5.2.4)$$

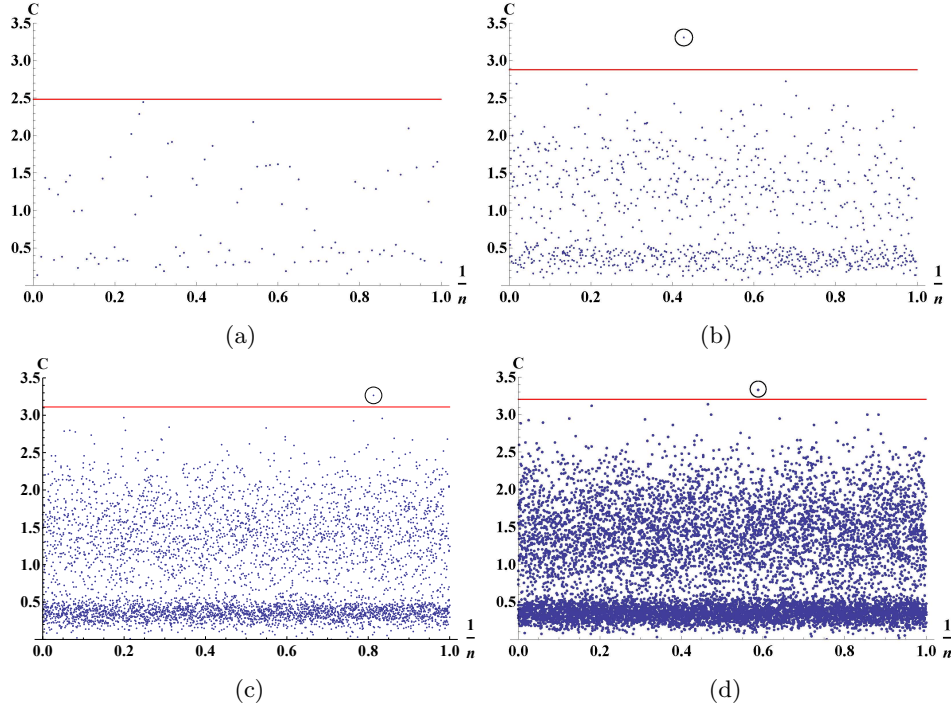


Figure 5.4: Point process simulation for the exceedance of the threshold $u = b_K$ in a time dependent environment where (a) $K=100$ (b) $K=1000$ (c) $K=5000$ (d) $K=10000$

Figure 5.4 gives simulation results for the users' capacities point process, and the exceedance above the threshold u as given in (5.2.4). One can see that as the population grows, the threshold value also grows in a way which maintains the average probability of exceedance $1/K$. We can see that in Sub-Figure 5.4a no exceedance occur among all users where in the other cases one user did manage to exceed alone.

5.2.3 Refinement On The Threshold Value

While considering *i.i.d.* random variables as the sequence of channel capacity, we can give a more accurate value for the threshold such that only one user exceeds it on average. The threshold given in the previous chapter (5.2.4), such that $u_n = b_K$, was formed due to the calculations of the mixed Gaussian distribution (5.1.1) which was the marginal distribution for a specific channel sample. As a reminder we are searching u_n such that

$$1 - F(u_n) = \frac{1}{K}$$

where in this chapter F is CDF of normal distribution with parameters (μ, σ) , which has a known complementary function in contrast to the mixed Gaussian distribution.

In [38] a use in the complimentary error function was done in order find threshold u_k such that k users will exceed it among the K users, in our case $k = 1$, hence we have

$$\begin{aligned} u_1 &= \mu + \sqrt{2} \sigma \operatorname{erfc}^{-1} \left(\frac{2}{K} \right) \\ &= \mu + \sigma \sqrt{2 \log(K) - \log \left[-2\pi \left(2 \log \left(\frac{1}{K} \right) + \log 2\pi \right) \right]} + o(K^2). \end{aligned} \quad (5.2.5)$$

Chapter 6

Performance analysis - distributed algorithm

In the previous chapter, we concerned ourselves with the question of what is the expected channel capacity in a distributed scheduling system based on the exceedance of some threshold u_K . This analysis touched just one part of the information communication chain, as we can ask relevant questions regarding delay and quality of service of the information, before it leaves the transmitter.

K users wish to transmit data to a shared access point. Users are not necessarily fully backlogged, and each of them has a Poisson arrival process with rate λ . Clearly this means a user may not have any information to send in a certain slot. We focus our attention on the users' queues, as illustrated in Figure 6.1, and, specifically, characterize the service process for each user. The users scheduling

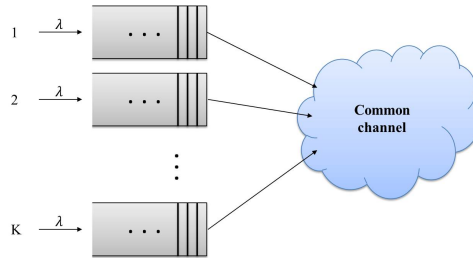


Figure 6.1: System model. K users access a common channel. Each user has a packet arrival process with rate λ

is based on the distributed algorithm, where a user checks if its channel capacity is above some threshold u_K , and transmits for one slot only if it exceeds it. The threshold is chosen such that *the probability of exceedance equals $1/K$* . Assume for now that this probability is a constant variable.

In the following, we investigate the behaviour of the system, starting with referring our queueing model to the known model of slotted aloha. We then review a known approximate model, which captures the system performance, and extend it to fit our model. Finally, we present two additional approaches, under different assumptions, to approximate this queueing system by the analogy to known types of queues. We do so by considering the process of threshold exceedance, which has Poisson properties and is asymptotically considered as a Poisson process.

In the first two subsections, we assume the users experience a time *independent channel*. Therefore, the rate for exceedance is the same for all users. In the last subsection we consider time dependency, and the possibility that users are in different channel states, hence, we have two different exceedance rates (as there is only value of the threshold).

6.1 Approximation by System and Users' State

It is not hard to see the resemblance of our model to the well explored slotted aloha system. Yet, the main difference is that in slotted aloha, the transmission scheme used is “immediate first transmission” i.e., if user i has an empty queue when a package arrives, it may transmit the package instantaneously, while in the case of collision, it may transmit again with a retransmission probability p_i . Our model follows a “delayed first transmission” scheme, where channel transmission can only happen with probability p_i (i.e., the threshold exceedance probability). When considering such a system consisting, of K queues, each with a memoryless arrival process yet a common server (i.e., the communication channel), the interdependence between the queues is very strong and cannot be ignored. It is reflected, among others, in the users' queue length, probabilities for successful transmission or collision and the delay of packages. Several works in the literature explored the behavior of such systems, with the same terms of QoS. Due to the interdependence between the queues, each of these works considered a different simplified mathematical model, which resulted with an approximation for the required metrics. We elaborate on one specific work [3], which introduced a relevant approximate model. Although this work introduced a general model for both “immediate” and “delayed” first transmission, the analytical derivations and numerical results refers only to the “immediate first transmission” scheme, and not the “delayed first transmission” scheme we are interested in. We will briefly review this model and emphasize the differences in the mathematical derivations due to the different schemes.

6.1.1 Review on the approximate model

The model presented in [3] assumed that the arrival rate at user i is λ_i , and the arrival processes are statistically independent between the users. Time is slotted and it takes exactly one slot to transmit one packet. Thus, λ_i is the probability of arrival for user i in any given slot. User i attempts, with probability p_i , to transmit the head-of-the-line packet in the queue (if the latter is nonempty) through a common channel to the receiver.

At the beginning of a given slot the users' status may be one of the three categories: *idle*, *active* or *blocked*. A user is idle if there are no packets in its queue at the end of the preceding slot. It is blocked if its queue is not empty and the latest attempted transmission was unsuccessful. It is active if its queue is not empty but its most recent attempted transmission was successful. We then let

$$p_i = \begin{cases} r_i \lambda_i & \text{if } i \text{ is idle,} \\ s_i & \text{if } i \text{ is active,} \\ q_i & \text{if } i \text{ is blocked.} \end{cases} \quad (6.1.1)$$

As can be seen, the setting of the problem is similar to the one considered in this study when $r_i = s_i = q_i = \frac{1}{K}$. As mentioned, in [3], the mathematical analysis and numerical results considered the case where $r_i = s_i = 1, q_i = p_i$, which is the original slotted aloha paradigm. Specifically, the model consist of a coupled Markov chains, the system-status chain and the queue-length chain. The transition probabilities of each chain, as will be presented later, depend on the steady-state probabilities of the other chain and thus, all state equations for both chains must be solved simultaneously. The system-status chain tries to capture the state of each user in any given time, hence, the status variable \bar{S} consists of K ternary variables, S_1, S_2, \dots, S_K , each of which indicates the status of the corresponding terminal. Namely,

$$S_i = \begin{cases} 0 & \text{if } i \text{ is idle,} \\ 1 & \text{if } i \text{ is active,} \\ 2 & \text{if } i \text{ is blocked.} \end{cases} \quad (6.1.2)$$

The total number of states achievable by this vector is given by

$$\sum_{i=0}^K \binom{K}{i} + K \sum_{i=0}^{K-1} \binom{K-1}{i} = \sum_{i=0}^K \binom{K}{i} i = 2^{K-1}(K+2) \quad (6.1.3)$$

where i indicates the number of blocked users, along with the observation which no more than one active user may be present in the system in any given time. The new calculations of the transition probabilities is according to the “delayed first transmission” scheme, which are different from the one presented in [3] displayed in appendix D. Two quantities which are needed for the calculation of the transition probabilities are:

$$\begin{aligned} P_i(1 | 1) &\triangleq P_r(\text{queue size} > 1 \mid \text{user } i \text{ is active}) \\ P_i(0 | 2) &\triangleq P_r(\text{queue size} = 1 \mid \text{user } i \text{ is blocked}). \end{aligned} \quad (6.1.4)$$

Note that these are not the transition probabilities of a specific chain, rather, they reflect the coupling between the two chains. Along with these quantities, one can calculate $P(\bar{S})$, the joint steady-state probability distribution of the random vector \bar{S} .

The queue-length Markov chain tracks both the status and queue length of each user, independently of the status and queue length of the other users. Specifically, the pair (T_i, N_i) represents the state of the user and $\pi(T_i, N_i)$ denotes its steady-state probability. N_i is the total number of packets at queue i and

$$T_i = \begin{cases} 0 & \text{if } i \text{ is blocked,} \\ 1 & \text{if } i \text{ is unblocked.} \end{cases}$$

The transition probabilities of the chain depend on the following average success probabilities:

$$\begin{aligned} P_B(i) &= P_r(\text{success} \mid \text{user } i \text{ is blocked}) \\ P_A(i) &= P_r(\text{success} \mid \text{user } i \text{ is active}) \\ P_I(i) &= P_r(\text{success} \mid \text{user } i \text{ is idle}), \end{aligned} \quad (6.1.5)$$

where the averaging is performed over the status of the other users (the status of the system). Namely, in order to calculate the probabilities (6.1.5) one needs the stationary distribution of the system-status chain $P(\bar{S})$. This leads to the coupling of the two sets of equations. The calculations of this average success probabilities are also different from the one presented in [3], and can be found in appendix E. The values which are in special interest to us, in order to help understand the system performance, are

$$\pi(0, 0) = \frac{\lambda_i \bar{P}_I(i)}{\lambda_i P_A(i) + \bar{\lambda}_i P_B(i)} \pi(1, 0) \quad (6.1.6)$$

$$\pi(1, 0) = \frac{\bar{\lambda}_i P_B(i) - \lambda_i \bar{P}_A(i)}{\bar{\lambda}_i P_B(i) - \lambda_i (P_I(i) - P_A(i))} \quad (6.1.7)$$

$$\pi(1, 1) = \frac{\lambda_i}{\bar{\lambda}_i} \pi(0, 0) \quad (6.1.8)$$

$$G_0^i(1) = \frac{\lambda_i \bar{\lambda}_i \bar{P}_I(i)}{\bar{\lambda}_i P_B(i) - \lambda_i (P_I(i) - P_A(i))} \quad (6.1.9)$$

$$G_1^i(1) = \lambda_i + \bar{\lambda}_i \frac{\bar{\lambda}_i P_B(i) - \lambda_i \bar{P}_A(i)}{\bar{\lambda}_i P_B(i) - \lambda_i (P_I(i) - P_A(i))} \quad (6.1.10)$$

where 6.1.9 and 6.1.10 are defined as the conditional moment generating function of the chain as follows

$$G_{T_i}^i(z) \triangleq \sum_{N_i=0}^{\infty} \pi(T_i, N_i) z^{N_i}, \quad T_i = 0, 1 \quad (6.1.11)$$

and $\bar{\lambda} = 1 - \lambda, \bar{P} = 1 - P$. The expressions presented here are the same ones as presented in [3], and also the mathematical development to attain them is the same. In order to find the steady state of the chains, as explained, the state equations must be solved simultaneously. The way to do so is by an iterative process, each time given the auxiliary quantities $P_i(1 | 1), P_i(0 | 2)$ and $P_I(i), P_A(i), P_B(i)$, as shown in Figure 6.2. The iterative method used is Wegstein’s iteration scheme, and after achieving satisfying convergence, one may calculate the performance metrics using the steady state of the chains. More details can be found in [3, 65].

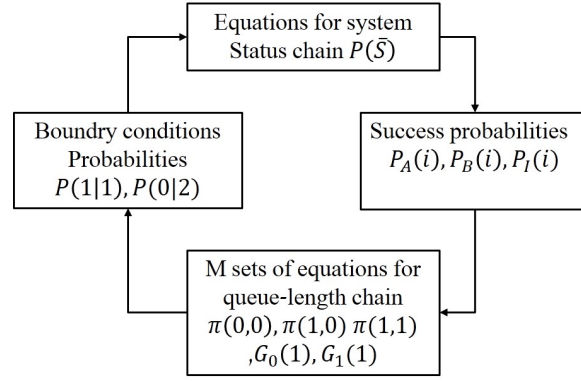


Figure 6.2: Iterative algorithm for the solution of the coupled sets of equations for the state probabilities of the system chain and the queue length chains.

The performance metrics

The delay a package endures from the moment it is generated and arrives to a user's queue until it is successfully transmitted, consists of three components, the time in line $W_q(i)$, the service time $W_s(i)$ and transmission time. Each of them is approximated separately and the sum of the three gives the total average delay. The following expression is the delay at a user i [3].

$$D_i = W_q(i) + W_s(i) + 1 \quad (6.1.12)$$

The 1 added represents the transmission time, which takes one slot. The service time is approximated by

$$W_s(i) = \frac{G_0(1)}{1 - \bar{\lambda}_i \pi(1, 0)} \cdot \frac{1}{P_B(i)}, \quad (6.1.13)$$

where the service time, as defined, is the time a package exists in the head of the queue. In case a successful transmission happens in the immediate arrival of a package, the service time is zero. In any other case, meaning the user is in blocked state and has packages in his queue, we counts the number of slots until a successful transmission, i.e., a geometric random variable. Note that in [3], the service time was approximated differently, considering idle slots as zero service time. Clearly, there is no meaning to service time with an empty queue, therefore, service time should be *normalized with the probability of not being in idle state with no arrival of a package*. The time in line is calculated by using Little's result, hence

$$W_q(i) = \frac{L_i}{\lambda_i} \quad (6.1.14)$$

where L_i is the average queue length of a user (without considering the blocked head-of-line packet as part of the queue), which is given by

$$L_i = \frac{\lambda_i^2 \bar{\lambda}_i \bar{P}_I(i)}{(\bar{\lambda}_i P_B(i) - \lambda_i \bar{P}_A(i))(\bar{\lambda}_i P_B(i) - \lambda_i (P_I(i) - P_A(i)))}. \quad (6.1.15)$$

The total weighted average system delay D is therefore

$$D = \frac{\sum_{i=1}^K D_i \lambda_i}{\sum_{i=1}^K \lambda_i}$$

where $\bar{\lambda} = 1 - \lambda$, $\bar{P} = 1 - P$. The probability for success given that the user is not idle, i.e., has package to send, and exceeds the threshold is given by

$$p_{succ}(i) = \frac{P_A(i)(G_1^i(1) - \pi(1, 0)) + P_I(i)\pi(1, 0) + P_B(i)G_0^i(1)}{p_i(1 - \pi(1, 0))} \quad (6.1.16)$$

where the numerator is the success probability regardless the users' state, divided by the probability that exceedance occur with a non empty queue.

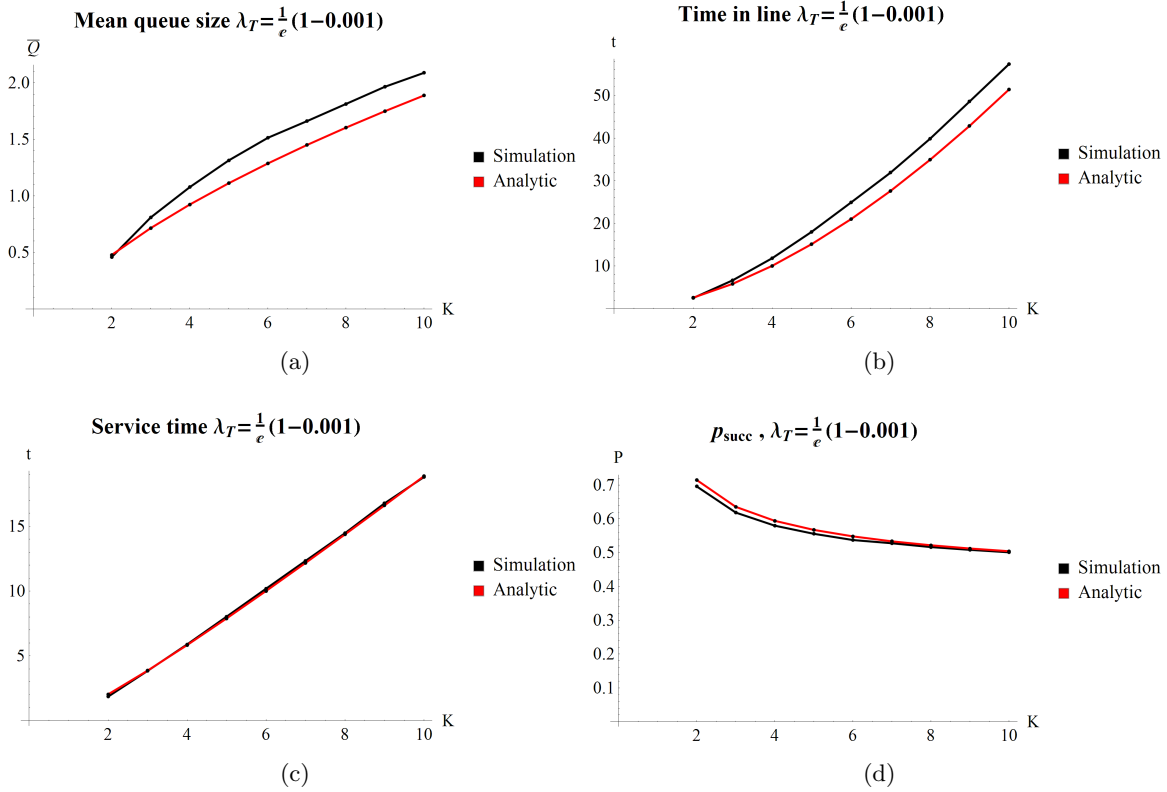


Figure 6.3: Analytic results and simulation for the system performance as a function of the number of users. Total arrival rate is $\lambda_T = \frac{1}{e}(1-0.001)$. (a) The mean queue size. (b) The time in line. (c) The service time. (d) The probability for success. The red line describes the analytic expressions (6.1.15), (6.1.14), (6.1.13), (6.1.16) respectively.

6.1.2 Performance analysis and comparisons

We present numerical results under “delayed first transmission” scheme. We divide the delay to its different components, i.e., the time in queue, service time, average queue size, system throughput and, in addition, the success probability p_{succ} given that a user has a package to send and exceeded the threshold. Due to the fact that there is no exact solution for the slotted aloha system, a comparison was made only with simulation results for this system. As explained earlier, the number of states in the system-status Markov chain grows exponentially with the number of users, therefore, the simulation and the numerical results were compared for 2 to 10 users. The performance parameters were also checked as a function of the arrival rates.

The simulation and analytic calculations were performed under the settings of a symmetric system. The arrival process and the threshold exceedance process were approximated by a Bernoulli process, due to the discrete nature of the system. Figure 6.3 depicts the performance metrics as a function of the number of users, where the total arrival rate is close to maximum value which still allows the system to be stable. The results depict very good match between the simulations and the analytic results given in (6.1.13), (6.1.14), (6.1.15) and (6.1.16).

Very important result is the fact that the probability for successful transmission, $p_{succ} = 1 - p_{coll}$, behaves exactly as predicted by the analytical model (Figure 6.3d), which means that it could be regarded as a *constant average probability regardless of the strong interdependence between the*

queues. This discovery is the foundation for the next two subsections, which, in contrast to the model considered thus far, are able to capture the system's behavior for large number of users.

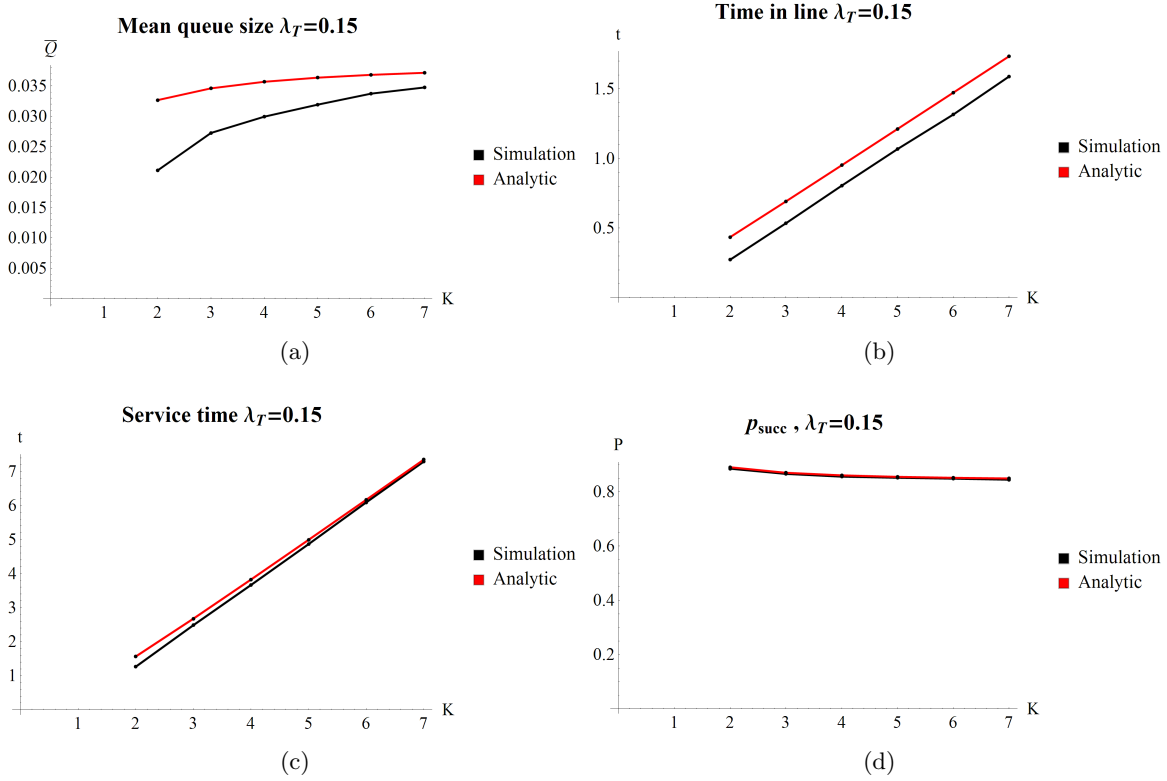


Figure 6.4: Analytic results and simulation for the system performance as a function of the number of users. Total arrival rate is $\lambda_T = 0.15$. (a) The mean queue size. (b) The time in line. (c) The service time. (d) The probability for success. The red line describes the analytic expressions (6.1.15), (6.1.14), (6.1.13), (6.1.16) respectively.

Figure 6.5 and 6.6 depicts the performance as a function of the total arrival rate, which help us deduce a few interesting observations. First, for small and moderate values of the arrival rate one can see very good agreement between the analytic results and the simulation while near the maximum value, the instability region, the values start to diverge, especially in Figures 6.5a and 6.5b. These results are consistent with the results of [3], which used small values of λ'_i 's as well. The second observation is that as the number of users grows, one can see even stronger compliance with the analytical results. This suggests that this approximate model may be suitable for a large population. Unfortunately, as mentioned, the calculation of the system steady state is intractable due to the exponential growth of the number of states with the number of users. In the following section, we give a simplified model, which utilizes the finding thus far and is able to approximate the performance for a large population as well.

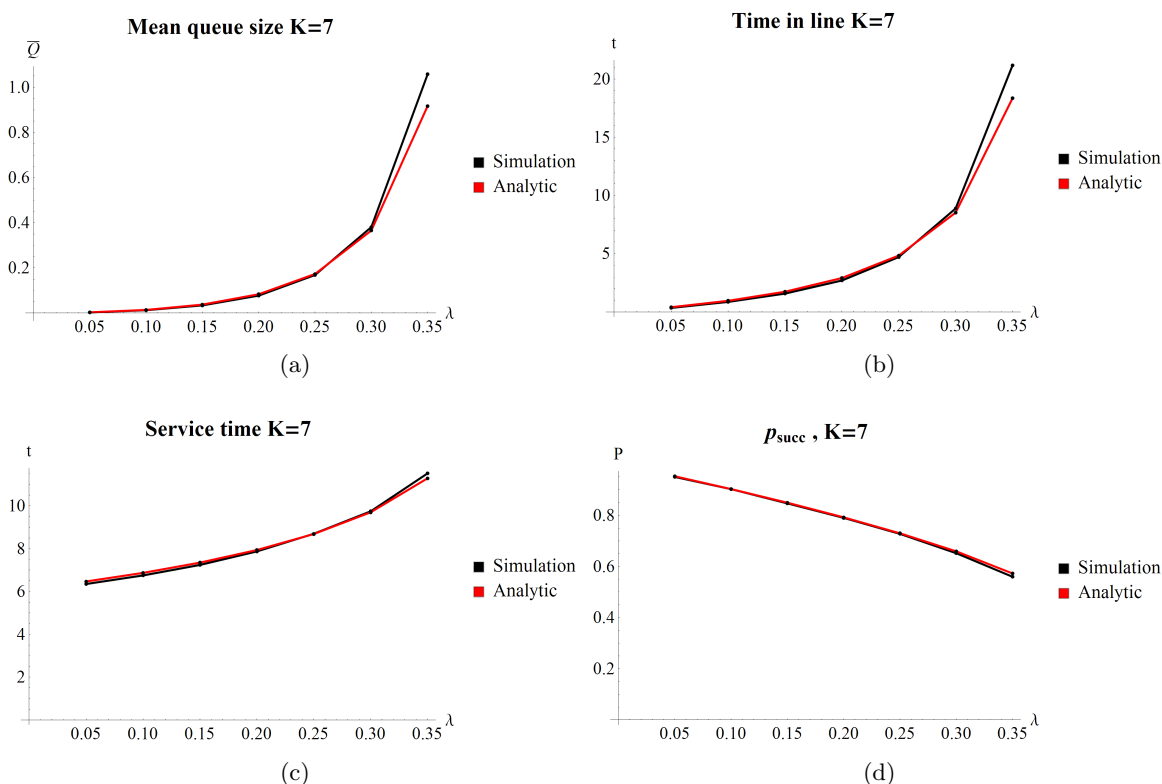


Figure 6.5: Analytic results and simulation for the system performance as a function of the total arrival rate. The number of users $K = 7$. (a) The mean queue size. (b) The time in line. (c) The service time. (d) The probability for success. The red line describes the analytic expressions (6.1.15), (6.1.14), (6.1.13), (6.1.16) respectively.

6.2 Approximation by Constant Collision Probability

From the previous section it is clear that for large population, the performance metrics of the system are hard to calculate. In this subsection we will describe the service time using a different approach, which assumes that the *probability of collision is constant* and equals the average collision probability. The legitimacy comes from the results given in the previous section. Another work which used a constant collision probability as the key approximation is [79]. There, a similar model was considered, with n users competing on a shared medium, however the users were *considered backlogged* and only throughput performance was investigated. Nevertheless the results of the approximation in [79] show good agreement with the simulations, which strengthens its applicability. In [66], independent $M/M/1$ queues were assumed, with a given distribution on the number of backlogged users. This results in an approximation with a varying collision probability. Herein, we use a different relaxation, assuming a *constant collision probability*. This allows us to easily decouple the queues. We start by investigating the exceedance process, by giving the terms for convergence to the Poisson process in the independent case of the channel, which connects directly to the service time. Then we turn to the service time analysis

6.2.1 Threshold exceedance process

The point process approximation as described in the preliminaries gives us convergence to a non-homogeneous Poisson process N of the points exceeding u_n . Here we shall use the notation given

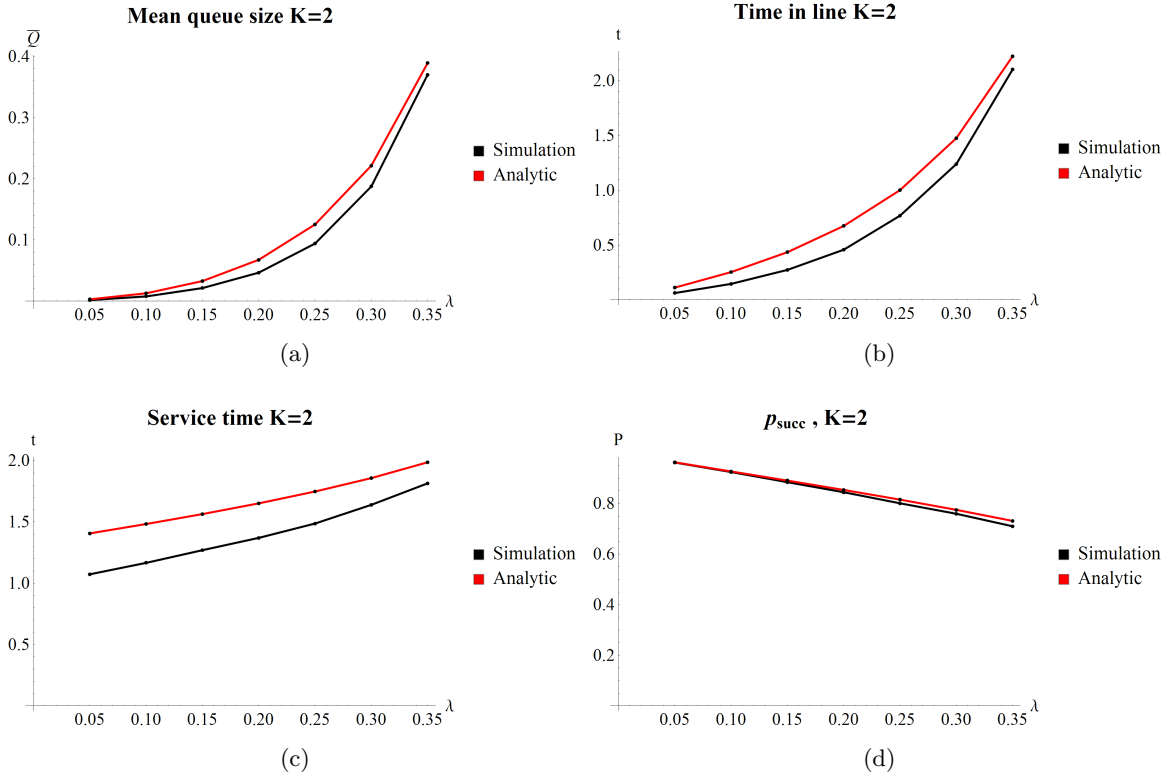


Figure 6.6: Analetic results and simulation for the system performance as a function of the total arrival rate. The number of users $K = 2$. (a) The mean queue size. (b) The time in line. (c) The service time. (d) The probability for success. The red line describes the analetic expressions (6.1.15), (6.1.14), (6.1.13), (6.1.16) respectively.

in [72] for the Poisson properties of exceedance.

Let C_1, C_2, \dots, C_n , a sequence of *i.i.d.* random variables with distribution F , be the channel capacity draws of a specific user. From [72, Theorem 2.1.1], if u_n satisfies $n(1 - F(u_n)) \rightarrow \tau$, then for $k = 0, 1, 2, \dots$,

$$P(N_n \leq k) \rightarrow e^{-\tau} \sum_{s=0}^k \frac{\tau^s}{s!}$$

where N_n is the number of exceedances of a level u_n by $\{C_n\}$. It is important to note that convergence in distribution happens if we change the “time scale” by a factor of n as defined in the definition of N_n in (3.4.1). Meaning that we have a point process in the interval $(0, 1]$ and the exceedance of such points has a limiting Poisson distribution. This restriction may be problematic in terms of the data arrival rate which will also need to be changed in a factor of n for the analysis. However, under certain conditions, there may be convergence of the exceeding points for the entire positive line in place of the unit interval. In [72, Theorem 5.2.1(ii)], it is stated that if for each $\tau > 0$, exists a sequence $\{u_n(\tau)\}$ satisfying $n(1 - F(u_n(\tau))) = nP(C_1 > u_n(\tau)) \rightarrow \tau$ as $n \rightarrow \infty$, and that $D(u_n(\tau)), D'(u_n(\tau))$ hold for all $\tau > 0$, then for any fixed τ , N_n converges in distribution to a Poisson process N on $(0, \infty)$ with parameter τ . Clearly the dependence conditions hold for this case, since the sequence $\{C_n\}$ is *i.i.d.* so we only need to show that the first condition holds for each τ .

Lemma 3. Assume F is the Gaussian distribution and let a_n and b_n be given according to (3.3.6)

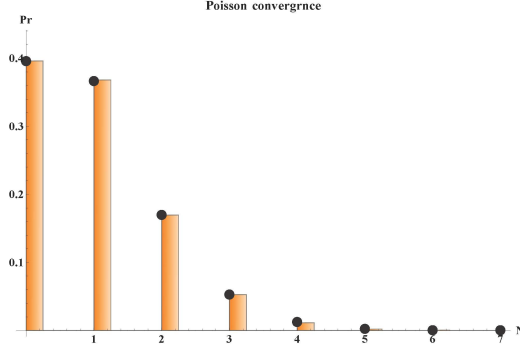


Figure 6.7: Simulation for 10000 user's capacities which follows an *i.i.d.* Gaussian distribution, showing the behaviour of exceedance which converge to a Poisson distribution (the dots) with parameter τ .

	Simulation	Poisson distribution
Pr(N=0)	0.3959	0.3961
Pr(N=1)	0.3682	0.3668
Pr(N=2)	0.1696	0.1698
Pr(N=3)	0.0527	0.0524
Pr(N=4)	0.0112	0.0121
Pr(N=5)	0.002	0.0022
Pr(N=6)	0.0004	0.0003
Pr(N=7)	0.0	0.00004

Figure 6.8: Values table of figure 6.7.

and (3.3.7) respectively. Fix any $\tau > 0$ and set $u_n(\tau) = \frac{\log 1/\tau}{a_n} + b_n$. Then,

$$\lim_{n \rightarrow \infty} n(1 - F(u_n(\tau))) = \tau.$$

Proof. In a similar way for the derivation of the normalizing constant in appendix A, and based on the derivation in [72, Theorem 1.5.3]. Let us find $u_n(\tau)$ which satisfies the equivalence condition for the convergence of the expression $n(1 - F(u_n(\tau)))$

$$\begin{aligned} n(1 - F(u_n(\tau))) &\rightarrow \tau \quad \text{as } n \rightarrow \infty \\ \frac{nf(u_n(\tau))}{u_n(\tau)} &\rightarrow \tau \quad \text{as } n \rightarrow \infty \end{aligned}$$

where the second line is true due to the Gaussian relation $1 - \Phi(u) \sim \frac{\phi(u)}{u}$ as $u \rightarrow \infty$, which in our case $u_n(\tau)$ grows with n . So

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} e^{-\frac{u_n^2(\tau)}{2}} &\rightarrow \frac{\tau u_n(\tau)}{n} \quad \text{as } n \rightarrow \infty \\ -\log \sqrt{2\pi} - \frac{u_n^2(\tau)}{2} &\rightarrow \log \tau + \log(u_n(\tau)) - \log n \quad \text{as } n \rightarrow \infty \end{aligned}$$

we know that $\log(u_n(\tau)) = \frac{1}{2}(\log 2 + \log \log n) + o(1)$, hence

$$\begin{aligned} \frac{u_n^2(\tau)}{2} &= \log \frac{1}{\tau} - \frac{1}{2} \log 4\pi - \frac{1}{2} \log \log n + \log n + o(1) \\ u_n^2(\tau) &= 2 \log n \left(1 + \frac{\log \frac{1}{\tau} - \frac{1}{2} \log 4\pi - \frac{1}{2} \log \log n}{\log n} + o\left(\frac{1}{\log n}\right) \right) \\ u_n(\tau) &= \sqrt{2 \log n} \left(1 + \frac{\log \frac{1}{\tau} - \frac{1}{2} \log 4\pi - \frac{1}{2} \log \log n}{2 \log n} + o\left(\frac{1}{\log n}\right) \right) \\ u_n(\tau) &= \frac{\log 1/\tau}{a_n} + b_n \end{aligned}$$

where the penultimate line is due to Taylor expansion. \square

After proving that all conditions hold and we have a convergence to a Poisson process on the real line of the exceeding points, we can conclude that a user tries to transmit at a rate of τ assuming he has packages to send. We will note this Poisson process of threshold exceedance as N_{exc} . Note that

the time between each user exceedance is distributed exponentially with parameter τ . The above leads us to a discussion on the duration of a slot.

At the limit of large number of users, the capacity seen by a user who exceeds the threshold is high (in fact, it scales like $O(\sigma\sqrt{2\log K} + \mu)$, see (5.1.8)). Hence, for any finite size of a data packet, transmission time tends to zero as the number of users increases. Furthermore, in this case, as convergence to Poisson process N_{exc} exists, events duration also goes to zero. This motivates us to suggest the asymptotic model, where transmission time is negligible but still exists, hence under this assumptions we will refer the slots as mini slots to emphasize this fact.

Remark 2 (Zero collisions). *The users are independent, and each of them is also characterized with the same N_{exc} process. As known, one of the properties of a Poisson process is that no two events can occur in the same time. Therefore, once a user exceeds u_n , assuming it's transmission time goes to zero, the probability that other users will exceed u_n goes to zero as well. While considering this scenario, we can observe an approximation for the behavior of the user's queue with Poisson arrival process with rate λ and exponential service process with parameter τ , as the familiar M/M/1 queue. This setting gives as an upper bound for the service time and other QoS properties. Still this scenario is treated only as a guide line for the analysis, since we are assuming that the services are considered to be synchronized, and hence collisions may occur and result with decline in the rate of service.*

6.2.2 Service time analysis

The service time, which is the time from the moment a package becomes first in queue, until it is successfully transmitted, depends on the rate which a user exceeds the threshold and the probability for success (the probability which collision did not occur), given that the user has a package to transmit. The time between threshold exceedances is exponentially distributed with parameter τ , due to the Poisson properties of the exceedance process (PPA). Therefore, the remaining parameter for the service time analysis is the probability for success. In fact, this probability is the main key for understanding this complex system. Due to the strong interdependence between the queues, this probability is constantly fluctuating and by quantifying it using average collision probability, one can approximate the service time.

We start with the assumption that the users are backlogged and give the appropriate conclusions regarding the success probability. Then we turn to our model with memoryless arrival process. From now on, we will refer to the probability of collision which is the complementary probability of success, i.e. $p_{coll} = 1 - p_{succ}$.

We define a collision scenario, if at least two users exceed the threshold u_n and both has packages to transmit. The analysis is done by considering a specific user i at the time which he exceeds u_n and assuming his queue, Q_i , is not empty. Given that, we note p_{coll} as the probability for collision,

$$p_{coll} = P\left(\bigcup_{j=1..K, j \neq i} C_j(n) > u_n, Q_j > 0 \mid C_i(n) > u_n, Q_i > 0\right) \quad (6.2.1)$$

where Q_j is the number of packages in the queue of user j . In general the probability p_{coll} depends on the states of the other users queues. We will analyse p_{coll} of a specific user, starting with the more simple analysis, resulting with an upper bound, assuming all users always have packages to transmit.

Backlogged system

In the case which all users are backlogged, the probability for collision is constat and independent among the users. The state of the users' queues does not change over time, since their queue is always full, and the only thing that impacts the probability for collision is whether a user exceeds u_n or not. Then, by observing the N_{exc} process, a certain point indicates a successful transmission with probability p_{succ} and a collision with probability $p_{coll} = 1 - p_{succ}$. Therefore a random

selection is made from N_{exc} resulting with a Poisson process of successful transmissions with rate $\tau \cdot (1 - p_{coll})$, which we will note as N_{tr} . The proof is quite obvious since the probability that an arrival occurs from the original process in the interval dt is τdt , which is independent of the arrivals outside the interval, where after the random selection, the probability for an arrival in the interval dt is $(1 - p_{coll}) \cdot \tau dt$ (independent of the arrivals outside the interval). Thus we conclude that the number of successful transmissions is a Poisson random variable with parameter $(1 - p_{coll}) \cdot \tau$. Nevertheless a well appointed proof can be found in appendix C.

Consider the following problem under the framework of the general problem. We have one user with arrival process with rate λ and $N_{tr}^{\text{backlogged}}$ process with rate $\tau \cdot (1 - p_{coll}^{\text{backlogged}})$, where all other users are backlogged. The definition of the problem in this manner will help us obtain an upper bound on p_{coll} by understanding the fact that the condition of a user in the original problem is better, in a way that users may have empty queues and although exceedance occur it will not result in a collision. This setting isolates for us the other users queues' states, from the analysis of the probability for collision.

In a given time the probability that a user exceeds the level u_n is $1 - F(u_n)$, so in order to have a collision, at least one user must exceed u_n among all users, hence we get

$$\begin{aligned} p_{coll}^{\text{backlogged}} &= P\left(\bigcup_{j=1..K, j \neq i} C_j(n) > u_n \mid C_i(n) > u_n, Q_i > 0\right) \\ &= 1 - P(C_1(n) < u_n, \dots, C_{K-1}(n) < u_n) \\ &= 1 - (F(u_n))^{K-1} = 1 - \left(1 - \frac{1}{K}\right)^{K-1} \\ &\xrightarrow{K \rightarrow \infty} 1 - e^{-1} \end{aligned} \tag{6.2.2}$$

Again the value u_n is set such that only one user on average exceeds it in each slot, hence $1 - F(u_n) = 1/K$.

This analysis is based on the fact that all the users have packages to send in any given time, so we can conclude that $p_{coll} \leq p_{coll}^{\text{backlogged}}$. Therefore a lower bound for the service process N_{tr} rate is $\tau \cdot (1 - p_{coll}^{\text{backlogged}})$.

Memoryless arrival

The assumption of constant and independent collision probability relaxes the analysis dramatically. We can now assume that each user's queue behaves like an $M/M/1$ queue, with a Poisson arrival process with rate λ and exponential service time with parameter $(1 - p_{coll}) \cdot \tau$, due to random selection with probability $1 - p_{coll}$ from the exceedance process. By doing so, we can use the known results for the probability of an empty $M/M/1$ queue, which is

$$P(Q = 0) = 1 - \rho = 1 - \frac{\lambda}{(1 - p_{coll}) \cdot \tau}, \tag{6.2.3}$$

where Q is the number of packages in the queue. We thus have the following lemma.

Lemma 4. *The users queue can be modeled as an $M/M/1$ queue with an arrival rate λ and service rate $(1 - p_{coll}) \cdot \tau$, where τ is the rate of the exceedance process, N_{exc} , and the probability for collision p_{coll} satisfies the equation*

$$p_{coll} = 1 - e^{-\frac{\lambda}{(1 - p_{coll}) \cdot \tau}}. \tag{6.2.4}$$

Proof. As random selection (with probability $(1 - p_{coll})$) is preformed on the exceedance Poisson process, this results in a Poisson process for successful transmission, which leads to an exponential

service time with parameter $(1 - p_{coll}) \cdot \tau$. Let us examine the probability p_{coll} . The probability that a specific user exceeds u_n , and its queue is not empty is

$$P(C_i > u_n, Q_i > 0) = P(C_i > u_n)P(Q_i > 0) = \frac{1}{K} \cdot \frac{\lambda}{\tau \cdot (1 - p_{coll})}.$$

Hence, given that one user exceeds u_n and its queue is not empty, p_{coll} is the probability that among all the other users which exceeded u_n , at least one has a package to send. Thus

$$\begin{aligned} p_{coll} &= \sum_{i=1}^{K-1} \binom{K-1}{i} \left(\frac{1}{K} \frac{\lambda}{\tau \cdot (1 - p_{coll})} \right)^i \left(1 - \frac{1}{K} \frac{\lambda}{\tau \cdot (1 - p_{coll})} \right)^{K-1-i} \\ &= 1 - \left(1 - \frac{1}{K} \frac{\lambda}{\tau \cdot (1 - p_{coll})} \right)^{K-1} \\ &\xrightarrow{K \rightarrow \infty} 1 - e^{-\frac{\lambda}{\tau \cdot (1 - p_{coll})}}. \end{aligned}$$

□

This simple result helps understanding the behaviour of the system's service time. Of course, equation (6.2.4) is an implicit equation and a numerical method is needed in order to find the value of p_{coll} . In Figure 6.9, one can see that the numerical calculation for the success probability (the blue line), as given in (6.2.4), coincides with the results of the simulation *and the approximate model from the previous section*. It is important to emphasise that for this comparison, the number of users is relatively small and therefore we used the equation in its explicit form, meaning without taking K to infinity. The good agreement of the results implies that for the case which $K \rightarrow \infty$, Lemma 4 is able to describe the behavior of the users' queues, hence the performance metrics of the system as a system of $M/M/1$ queues. Simulations for large number of users show excellent agreement as well, as can be seen in Figure 6.10, where a comparison of the average service time was made between an $M/M/1$ queue with service rate $p_{succ}\tau$, which was calculated according to equation (6.2.4), with simulation of the system, which of course includes dependent queues and variable p_{succ} .

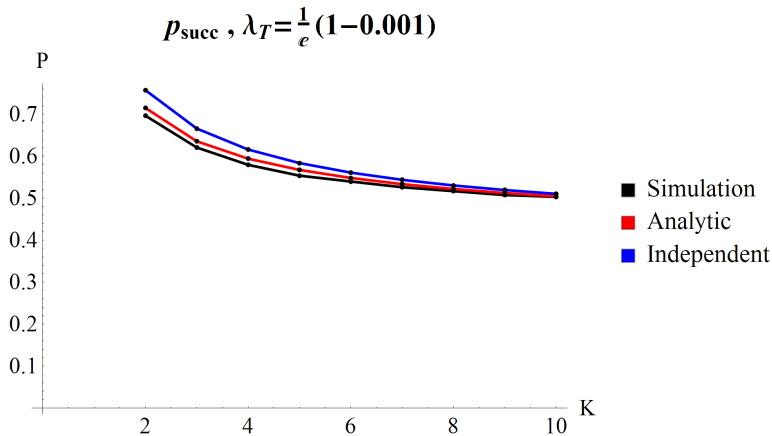


Figure 6.9: Comparison of the success probability between the analytic derivations of the approximate model I, simulation results, and the approximation of K independent $M/M/1$ queues given in Lemma 4.

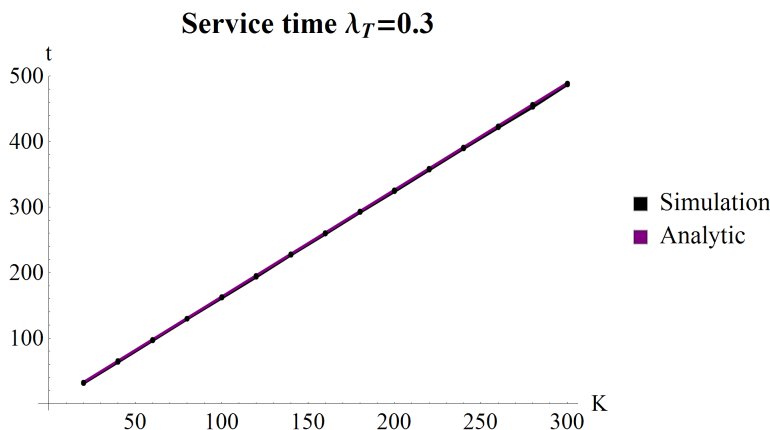


Figure 6.10: The service time of an M/M/1 queue with service rate $p_{succ}\tau$, compared to simulation results of system with K interdependent queues.

6.3 Approximation by Constant Collision Probability - Time Dependent Channel

In chapter 5 we analyzed the channel capacity, where each user experiences a time dependent process for the channel gain. There we assumed that all users are backlogged, therefore queuing aspects were not taken into account. We showed that under our distributed scheme the exceedance process of a specific user, above some threshold u_K , is exponentially distributed with some rate τ . This rate of exceedance was derived from the stationary distribution of the gain process, and there is no real separation between users that are in good state or in bad state in means of different rates for the corresponding states. The difference is embodied in the stationary distribution. thus τ is the average rate of exceedance for such a system, which a user spends some of the time in the good state and some in the bad state, according to the transition probabilities.

Then, in the beginning of chapter 6 we turned to the analysis of the queues in such distributed multiple-access system, when we assumed a memoryless arrival process. In section 6.1 we gave an elaboration for a known approximate model which gives us performance parameters for such system, and due to its bad scalability for a large population we gave in section 6.2 another approach to approximate the users' queues in a large population system. We show that the collision probability in each slot may be treated as constant, where this is called decoupling approximation, which works nicely when the number of users is large. Both sections' analysis was preformed while assuming that the users' capacities are *i.i.d.* random variables and therefore the threshold value was set accordingly.

It's an interesting question, considering the stationary good-bad capacity sequence as described earlier in (4.1.1), is it enough to set only the threshold value for the stationary case, as given in 5.2.4, with the relevant normalizing constants, as given in Theorem 8, to capture the service time behaviour as described in Lemma 4. We will not try to answer this question now, but rather propose another model which combines the queuing system with the users' good-bad channel process.

Assume now that each user experiences a good or bad channel, according to the time dependent model in Figure 4.1. Assume further that the transitions between the states are exponentially distributed with rates α and β . Depending on the state a user exists in, it exceeds the threshold with a different rate. Namely, if a user is in a good state, it would exceed the threshold more often than when being in a bad state. Hence, the user's queue is modeled as follows. A Poisson arrival process with rate λ , and an exceedance process which is assumed to be exponentially distributed according to the user's state. If a user is in a good state, the exceedance process is exponentially distributed

with rate μ_g , and if it is in a bad state, the exceedance process is exponentially distributed with rate μ_b .

Since we consider multiple user system, where each user competes for the channel distributively, collisions may occur. Since we consider this probability as constant, we have a similar conclusion, that is, random selection is performed from the exceedance process, with probability p_{succ} , regardless the state which the user exists in. This results in an exponentially distributed service process with rates $\mu_g \cdot p_{succ}$ or $\mu_b \cdot p_{succ}$, depending on the good-bad state respectively. The individual queue model for this case is presented in Figure 6.11.

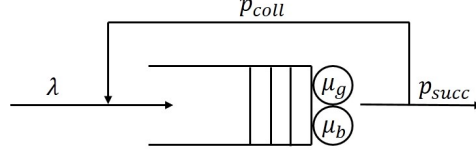


Figure 6.11: Queue model with varying service rates according to a time dependent channel.

This time-dependent queue model can be represented as a continuous Markov process on the set of states $\{\pi_m^i\}$ for $i \in \{b, g\}$, which indicates the good or bad state, and $m = 0, 1, 2, \dots$ the number of packages in the queue. This two dimensional Markov chain is presented in Figure 6.12. To ease notation, we denote $\mu'_g = \mu_g \cdot p_{succ}$ and $\mu'_b = \mu_b \cdot p_{succ}$.

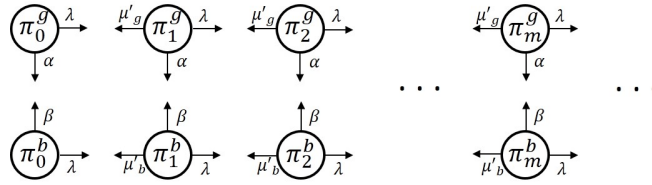


Figure 6.12: Queue model for time dependent user

In [80], the authors investigated a problem with heterogeneous arrivals and service to an $M/M/1$ queue, which was solved by analysing a similar Markov chain to the one presented in Figure 6.12. We make use of the solution in [80], with the right adjustment to our model. The system is solved while using generating-function techniques, resulting with a solution of the steady state probabilities, $\{\pi_m^i\}$, as a function of the transition rate parameters and the root of a third degree polynomial $g(z)$ (only one exist under the assumptions). Here, we will show the relevant fitted steady state results of the chain. First, let us define

$$\hat{\mu} = \pi_g \mu'_g + \pi_b \mu'_b$$

where

$$\pi_g = \frac{\beta}{\alpha + \beta}$$

$$\pi_b = \frac{\alpha}{\alpha + \beta}$$

$\hat{\mu}$ is the average service rate. It is important to emphasize that in order to maintain stability of the system, we require that $\hat{\mu} > \lambda$. We define the partial generating functions of the system as

$$G_i(z) = \sum_{m=0}^{\infty} \pi_m^i z^m \quad | \quad z| \leq 1, \quad i = g, b.$$

which eventually equals to

$$G_g(z) = (\beta(\hat{\mu} - \lambda)z + \pi_0^g \mu'_g (1 - z)(\lambda z - \mu'_b)) / g(z)$$

$$G_b(z) = (\alpha(\hat{\mu} - \lambda)z + \pi_0^b \mu'_b (1 - z)(\lambda z - \mu'_g)) / g(z),$$

where $g(z)$ is

$$g(z) = \lambda^2 z^3 - (\alpha\lambda + \beta\lambda + \lambda^2 + \lambda\mu'_b + \lambda\mu'_g)z^2 + (\alpha\mu'_b + \beta\mu'_g + \mu'_g\mu'_b + \lambda\mu'_b + \lambda\mu'_g)z - \mu'_g\mu'_b \quad (6.3.1)$$

The steady state probabilities for an empty queue, depending on the system state, are

$$\begin{aligned} \pi_0^g &= \frac{\beta(\hat{\mu} - \lambda)z_0}{\mu'_g(1 - z_0)(\mu'_b - \lambda z_0)} \\ \pi_0^b &= \frac{\alpha(\hat{\mu} - \lambda)z_0}{\mu'_b(1 - z_0)(\mu'_g - \lambda z_0)} \end{aligned} \quad (6.3.2)$$

where z_0 is the root of the polynomial $g(z)$. The remaining steady state probabilities are as follows

$$\begin{aligned} \pi_m^g &= \pi_{m-1}^g \frac{\lambda}{\mu'_g} + \sum_{j=0}^{m-1} \pi_j^g \frac{\alpha}{\mu'_g} - \sum_{j=0}^{m-1} \pi_j^b \frac{\beta}{\mu'_g} \quad m > 0 \\ \pi_m^b &= \pi_{m-1}^b \frac{\lambda}{\mu'_b} + \sum_{j=0}^{m-1} \pi_j^b \frac{\beta}{\mu'_b} - \sum_{j=0}^{m-1} \pi_j^g \frac{\alpha}{\mu'_b} \quad m > 0 \end{aligned} \quad (6.3.3)$$

From the first derivative of the partial generating functions, we can attain the expected queue size, which includes the head of line package, unlike the analysis in subsection 6.1

$$\bar{Q} = \frac{\lambda}{\hat{\mu} - \lambda} + \frac{\mu'_g(\mu'_b - \lambda)\pi_0^g + \mu'_b(\mu'_g - \lambda)\pi_0^b - (\mu'_g - \lambda)(\mu'_b - \lambda)}{(\alpha + \beta)(\hat{\mu} - \lambda)}$$

Therefore, using Little's theorem we can attain the average waiting time in the queue:

$$W = \frac{\bar{Q}}{\lambda}.$$

For the completeness of the above results, we need to calculate p_{succ} , which until now is a fixed and unknown parameter. Hence, in the same manner as the previous subsection, we define the probability that a specific user exceeds the threshold and its queue is not empty, to be

$$\begin{aligned} P_t &\triangleq Pr(\text{transmission attempt}) \\ &= Pr(\text{exceedance occur \& queue is not empty}) \\ &= (G_g(1) - \pi_0^g)(1 - e^{-\mu_g}) + (G_b(1) - \pi_0^b)(1 - e^{-\mu_b}), \end{aligned} \quad (6.3.4)$$

where the first term consists of the probability to be in a good state with a package to transmit, times the probability which an exceedance occurs while existing in the good state; The second term is similar, referring only to the bad state. The probability for success, given one user is about to transmit, can be obtained by a similar calculation as in Lemma 4, and we have

$$p_{succ} = (1 - P_t)^{K-1}. \quad (6.3.5)$$

Now, in order to attain this probability, the above equation and the third degree polynomial (6.3.1) must be solved simultaneously, since the root z_0 and p_{succ} are coupled. As $K \rightarrow \infty$, the probability p_{succ} remains constant and does not go to zero since the probability for a transmission attempt, P_t , scales like $O(1/K)$. This is true since in the first place, the threshold value was set such that on average, the probability for exceedance equals $1/K$. Thus, $(1 - e^{-\mu_g}) = O(c_g/K)$ and $(1 - e^{-\mu_b}) = O(c_b/K)$, where c_g, c_b are some constants which satisfy $c_g > c_b$. Letting $K \rightarrow \infty$, we get a constant probability. In Figure 6.13 one can see the probability for success p_{succ} , which coincides with the simulation results and therefore the good agreement of the system performance with the simulations, as presented in Figure 6.14. The simulations were performed while assuming equal transitions rates between the good and bad state, i.e. $\alpha = \beta = 0.1$. Since we are considering

a symmetric system, the arrival and service rate parameters, as mentioned in the figures, were considered as the total rates of the system and were divided equally among all the users.

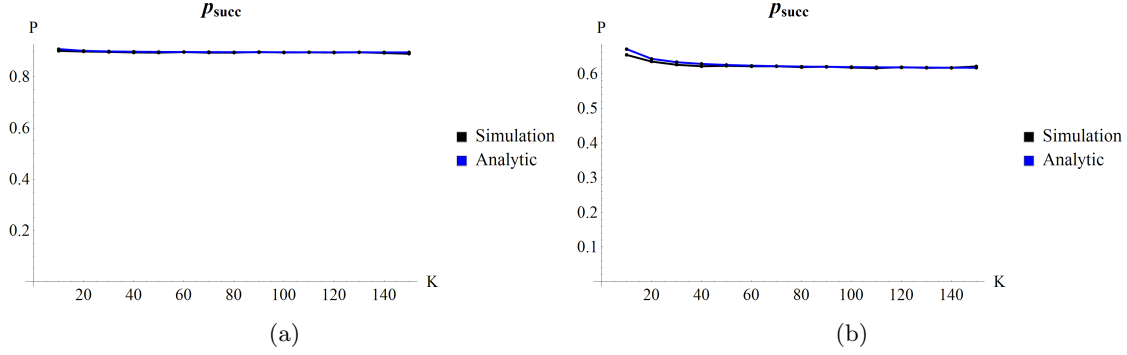


Figure 6.13: Queueing system success probability results as a function of the number of users in a time dependent queue model. $\lambda_T = \{0.1, 0.3\}$, $\mu_g = 0.7$, $\mu_b = 0.5$

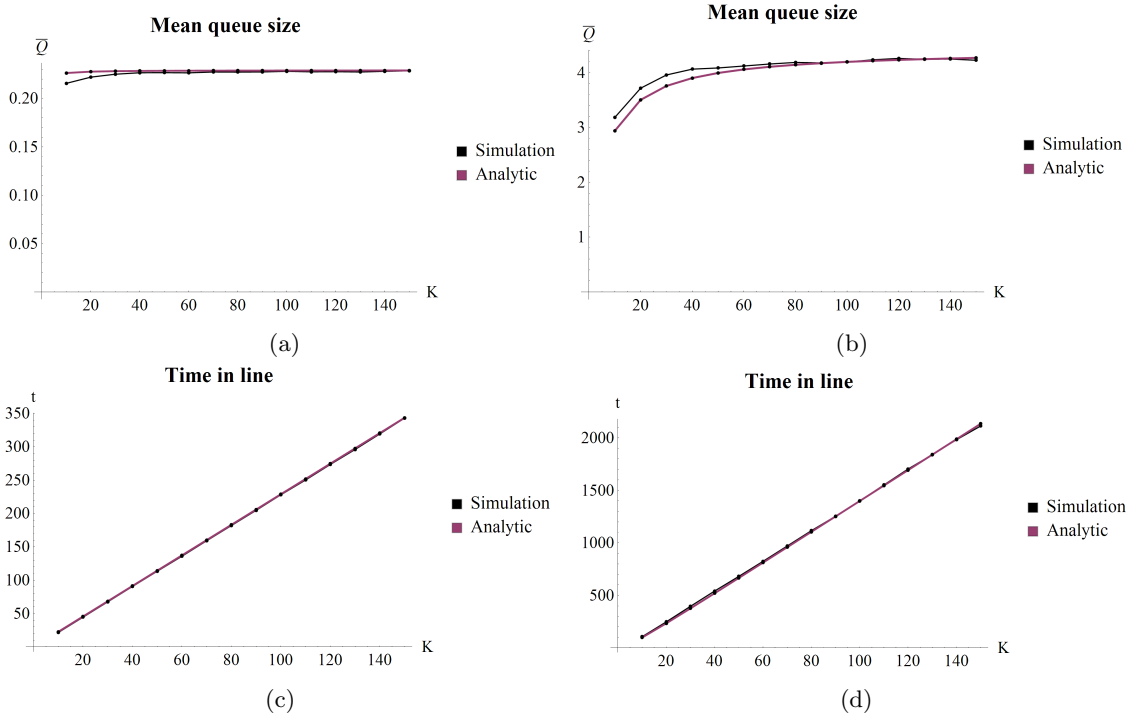


Figure 6.14: Queueing system performance results as a function of the number of users in a time dependent queue model. $\lambda_T = \{0.1, 0.3\}$, $\mu_g = 0.7$, $\mu_b = 0.5$

Chapter 7

Conclusions

This work investigated the channel capacity and the performance of a multi-user MAC system in a time dependent environment under both centralized and distributed scheduling. Specifically, the expected channel capacity gain was derived in the case which the dependent capacity sequence was modeled as a stationary process characterized by the good bad channel Markov process. If the long range dependence condition $D(u_n)$ and the local dependence condition $D'(u_n)$ hold, the extreme laws are the same as if the sequence was *i.i.d.* but with different normalizing constants. A distributed scheduling algorithm for this system was suggested as well, and we showed that both approaches have the same scaling laws, the distributed approach being smaller only by a factor of e^{-1} , so there was no loss of optimality due to the distributed algorithm.

The performance of the system was derived while considering queueing theory aspects. In fact this precise characterization is a very difficult mission, which up until today was not solved. Therefore we presented approximation models to describe its behaviour. First we address to the *i.i.d.* case where the users do not experience a time varying channel. For that case we elaborated an existing model and showed results for our paradigm. In addition we gave another approach, which relates the queues as independent, concerning the probability of collision in the random access mechanism, and enables us to consider them each as a much more simple queue. Lastly we tied the capacity part of this work with the queueing part when we suggested a queue model, which is time dependent and modeled by our good-bad channel model. We showed good agreement between the analytic model and the simulation results.

Appendix A

Convergence to Extreme Value Distribution

A.1 Convergence proof

Proof. In the following proof we give compliance for the convergence conditions and the derivation of a_K and b_K . The stationary distribution $F(t) = pF_g(t) + qF_b(t)$ as shown earlier has a negative derivative f' from $x_0 = \max\{\mu_g, \mu_b\}$ till ∞ . So we only need to show that (5.1.2) holds,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{f'(t)(1 - F(t))}{f^2(t)} &= \lim_{t \rightarrow \infty} \frac{(pf'_g(t) + qf'_b(t))(1 - (pF_g(t) + qF_b(t)))}{(pf_g(t) + qf_b(t))^2} \\ &= \lim_{t \rightarrow \infty} \frac{\frac{1}{2}(pf'_g(t) + qf'_b(t)) \left(p \operatorname{Erfc} \left(\frac{t - \mu_g}{\sqrt{2}\sigma_g} \right) + q \operatorname{Erfc} \left(\frac{t - \mu_b}{\sqrt{2}\sigma_b} \right) \right)}{(pf_g(t) + qf_b(t))^2} \end{aligned}$$

In [81], 7.1.13, we can find upper and lower bounds for the complementary error function,

$$\frac{2}{\sqrt{\pi}} \frac{e^{-t^2}}{t + \sqrt{t^2 + 2}} < \operatorname{Erfc}(t) \leq \frac{2}{\sqrt{\pi}} \frac{e^{-t^2}}{t + \sqrt{t^2 + \frac{4}{\pi}}}.$$

where these inequalities are true for $t > 0$ which fits our case for $t \rightarrow \infty$. Using these bounds we will show with the sandwich rule that the limit above converge to -1 . Let us consider first the lower bound of the complementary error function,

$$\begin{aligned} &\lim_{t \rightarrow \infty} \frac{\frac{1}{2}(pf'_g(t) + qf'_b(t)) \left(p \frac{2}{\sqrt{\pi}} \frac{e^{-\frac{(t - \mu_g)^2}{2\sigma_g^2}}}{\frac{t - \mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t - \mu_g)^2}{2\sigma_g^2} + 2}} + q \frac{2}{\sqrt{\pi}} \frac{e^{-\frac{(t - \mu_b)^2}{2\sigma_b^2}}}{\frac{t - \mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t - \mu_b)^2}{2\sigma_b^2} + 2}} \right)}{(pf_g(t) + qf_b(t))^2} \\ &= \lim_{t \rightarrow \infty} \frac{1}{\sqrt{\pi}} \frac{(pf'_g(t) + qf'_b(t)) \left(p \frac{\sqrt{2\pi}\sigma_g f_g(t)}{\frac{t - \mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t - \mu_g)^2}{2\sigma_g^2} + 2}} + q \frac{\sqrt{2\pi}\sigma_b f_b(t)}{\frac{t - \mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t - \mu_b)^2}{2\sigma_b^2} + 2}} \right)}{(pf_g(t) + qf_b(t))^2} \\ &= \lim_{t \rightarrow \infty} -\sqrt{2} \frac{\left(pf_g(t) \frac{t - \mu_g}{\sigma_g^2} + qf_b(t) \frac{t - \mu_b}{\sigma_b^2} \right) \left(p \frac{\sigma_g f_g(t)}{\frac{t - \mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t - \mu_g)^2}{2\sigma_g^2} + 2}} + q \frac{\sigma_b f_b(t)}{\frac{t - \mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t - \mu_b)^2}{2\sigma_b^2} + 2}} \right)}{(pf_g(t) + qf_b(t))^2} \end{aligned}$$

The Limit above can be break to four different limits,

$$\begin{aligned} & \lim_{t \rightarrow \infty} -\sqrt{2}p^2 \frac{f_g(t)^{\frac{t-\mu_g}{\sigma_g^2}} \frac{\sigma_g f_g(t)}{\frac{t-\mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t-\mu_g)^2}{2\sigma_g^2} + 2}}}{(pf_g(t) + qf_b(t))^2} + \lim_{t \rightarrow \infty} -\sqrt{2}q^2 \frac{f_b(t)^{\frac{t-\mu_b}{\sigma_b^2}} \frac{\sigma_b f_b(t)}{\frac{t-\mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t-\mu_b)^2}{2\sigma_b^2} + 2}}}{(pf_g(t) + qf_b(t))^2} + \\ & \lim_{t \rightarrow \infty} -\sqrt{2}pq \frac{f_g(t)^{\frac{t-\mu_g}{\sigma_g^2}} \frac{\sigma_b f_b(t)}{\frac{t-\mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t-\mu_b)^2}{2\sigma_b^2} + 2}}}{(pf_g(t) + qf_b(t))^2} + \lim_{t \rightarrow \infty} -\sqrt{2}pq \frac{f_b(t)^{\frac{t-\mu_b}{\sigma_b^2}} \frac{\sigma_g f_g(t)}{\frac{t-\mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t-\mu_g)^2}{2\sigma_g^2} + 2}}}{(pf_g(t) + qf_b(t))^2} \end{aligned}$$

please notice that the first and the second limits are similar with the exception of their indexes, and so does the third and the fourth limit. We start with the first limit calculation,

$$\begin{aligned} & \lim_{t \rightarrow \infty} -\sqrt{2}p^2 \frac{f_g(t)^{\frac{t-\mu_g}{\sigma_g^2}} \frac{\sigma_g f_g(t)}{\frac{t-\mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t-\mu_g)^2}{2\sigma_g^2} + 2}}}{(pf_g(t) + qf_b(t))^2} = \lim_{t \rightarrow \infty} -\sqrt{2} \frac{p^2 f_g^2(t) (t - \mu_g) \sigma_g}{(pf_g(t) + qf_b(t))^2 \sigma_g^2 \frac{t-\mu_g}{\sqrt{2}\sigma_g} \left(1 + \sqrt{1 + \frac{4\sigma_g^2}{(t-\mu_g)^2}}\right)} \\ & = \lim_{t \rightarrow \infty} -2 \frac{p^2 f_g^2(t)}{(pf_g(t) + qf_b(t))^2 \left(1 + \sqrt{1 + \frac{4\sigma_g^2}{(t-\mu_g)^2}}\right)} = \lim_{t \rightarrow \infty} -2 \frac{p^2 f_g^2(t)}{(pf_g(t) + qf_b(t))^2} \cdot \lim_{t \rightarrow \infty} \frac{1}{\left(1 + \sqrt{1 + \frac{4\sigma_g^2}{(t-\mu_g)^2}}\right)} \\ & = \lim_{t \rightarrow \infty} -2 \frac{p^2 f_g^2(t)}{(pf_g(t) + qf_b(t))^2} \cdot \frac{1}{2} = \lim_{t \rightarrow \infty} - \frac{p^2 f_g^2(t)}{(pf_g(t) + qf_b(t))^2} = - \left(\lim_{t \rightarrow \infty} \frac{pf_g(t)}{(pf_g(t) + qf_b(t))} \right)^2 \\ & \stackrel{(a)}{=} - \left(\lim_{t \rightarrow \infty} \frac{1}{1 + \frac{qf_b(t)}{pf_g(t)}} \right)^2 = - \left(\frac{1}{1 + \lim_{t \rightarrow \infty} \frac{qf_b(t)}{pf_g(t)}} \right)^2 = \\ & \Rightarrow \lim_{t \rightarrow \infty} \frac{qf_b(t)}{pf_g(t)} = \frac{\sigma_g}{\sigma_b} \frac{q}{p} \lim_{t \rightarrow \infty} e^{-\frac{(t-\mu_b)^2}{2\sigma_b^2} + \frac{(t-\mu_g)^2}{2\sigma_g^2}} = \frac{\sigma_g}{\sigma_b} \frac{q}{p} e^{\lim_{t \rightarrow \infty} -\frac{(t-\mu_b)^2}{2\sigma_b^2} + \frac{(t-\mu_g)^2}{2\sigma_g^2}} \\ & \Rightarrow \lim_{t \rightarrow \infty} \frac{t^2(\sigma_b^2 - \sigma_g^2) + t(2\mu_b\sigma_g^2 - 2\mu_g\sigma_b^2) + C}{2\sigma_b^2\sigma_g^2} = \begin{cases} \infty & \sigma_g^2 < \sigma_b^2 \\ -\infty & \sigma_g^2 \geq \sigma_b^2 \end{cases} \text{ assuming } \mu_g > \mu_b \end{aligned}$$

So,

$$\lim_{t \rightarrow \infty} \frac{qf_b(t)}{pf_g(t)} = \begin{cases} \infty & \sigma_g^2 < \sigma_b^2 \\ 0 & \sigma_g^2 \geq \sigma_b^2 \end{cases} \text{ assuming } \mu_g > \mu_b$$

Note that in (a) we assume that $p \neq 0$. This assumption implies that the situation which all the users are in the bad group is not taken in consideration. For that case all the users have the same channel and the analysis is known and not in our interest. In the same manner we assume also that $q \neq 0$ for the opposite situation. Hence, the first limit result is

$$\lim_{t \rightarrow \infty} -\sqrt{2}p^2 \frac{f_g(t)^{\frac{t-\mu_g}{\sigma_g^2}} \frac{\sigma_g f_g(t)}{\frac{t-\mu_g}{\sqrt{2}\sigma_g} + \sqrt{\frac{(t-\mu_g)^2}{2\sigma_g^2} + 2}}}{(pf_g(t) + qf_b(t))^2} = \begin{cases} 0 & \sigma_g^2 < \sigma_b^2 \\ -1 & \sigma_g^2 \geq \sigma_b^2 \end{cases} \text{ assuming } \mu_g > \mu_b$$

As mentioned earlier the first and the second limits different only in their indexes, therefore the result for the second limit is

$$\lim_{t \rightarrow \infty} -\sqrt{2}q^2 \frac{f_b(t)^{\frac{t-\mu_b}{\sigma_b^2}} \frac{\sigma_b f_b(t)}{\frac{t-\mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t-\mu_b)^2}{2\sigma_b^2} + 2}}}{(pf_g(t) + qf_b(t))^2} = \begin{cases} -1 & \sigma_g^2 < \sigma_b^2 \\ 0 & \sigma_g^2 \geq \sigma_b^2 \end{cases} \text{ assuming } \mu_g > \mu_b$$

We turn now for the third limit calculation

$$\begin{aligned}
& \lim_{t \rightarrow \infty} -\sqrt{2}pq \frac{f_g(t) \frac{t-\mu_b}{\sigma_g^2} \frac{\sigma_b f_b(t)}{\frac{t-\mu_b}{\sqrt{2}\sigma_b} + \sqrt{\frac{(t-\mu_b)^2}{2\sigma_b^2} + 2}}}{(pf_g(t) + qf_b(t))^2} = \lim_{t \rightarrow \infty} -\sqrt{2} \frac{pqf_g(t) (t-\mu_g)\sigma_b f_b(t)\sqrt{2}\sigma_b}{(pf_g(t) + qf_b(t))^2 \sigma_g^2 (t-\mu_b) \left(1 + \sqrt{1 + \frac{4\sigma_b^2}{(t-\mu_b)^2}}\right)} \\
& = -2 \frac{\sigma_b^2}{\sigma_g^2} \lim_{t \rightarrow \infty} \frac{pf_g(t)}{(pf_g(t) + qf_b(t))} \cdot \lim_{t \rightarrow \infty} \frac{qf_b(t)}{(pf_g(t) + qf_b(t))} \cdot \lim_{t \rightarrow \infty} \frac{(t-\mu_g)}{(t-\mu_b)} \cdot \lim_{t \rightarrow \infty} \frac{1}{\left(1 + \sqrt{1 + \frac{4\sigma_g^2}{(t-\mu_g)^2}}\right)} \\
& = -\frac{\sigma_b^2}{\sigma_g^2} \lim_{t \rightarrow \infty} \frac{pf_g(t)}{(pf_g(t) + qf_b(t))} \cdot \lim_{t \rightarrow \infty} \frac{qf_b(t)}{(pf_g(t) + qf_b(t))} \\
& = -\frac{\sigma_b^2}{\sigma_g^2} \begin{cases} 0 & \sigma_g^2 < \sigma_b^2 \\ 1 & \sigma_g^2 \geq \sigma_b^2 \end{cases} \text{ assuming } \mu_g > \mu_b \cdot \begin{cases} 1 & \sigma_g^2 < \sigma_b^2 \\ 0 & \sigma_g^2 \geq \sigma_b^2 \end{cases} \text{ assuming } \mu_g > \mu_b = 0
\end{aligned}$$

The fourth limit also share the same result, and if we add all the parts we can see that the limit converge to -1 . If we consider the upper bound of the complementary error function we will find also that the limit convergence to -1 since the analytical development is the same and the limit

$$\lim_{t \rightarrow \infty} \frac{1}{\left(1 + \sqrt{1 + \frac{4\sigma_g^2}{(t-\mu_g)^2}}\right)} = \lim_{t \rightarrow \infty} \frac{1}{\left(1 + \sqrt{1 + \frac{8\sigma_g^2}{\pi(t-\mu_g)^2}}\right)} = \frac{1}{2}$$

Therefore we can conclude that condition (5.1.2) holds for our stationary distribution. \square

We now show that the second condition (5.1.3) also holds for the stationary distribution. Let us examine the expression:

$$1 - F(t) = p(1 - F_g(t)) + q(1 - F_b(t))$$

Since $F_i(t)$ is a Gaussian distribution we shall use the asymptotic relation:

$$1 - \Phi(t) \sim \frac{\phi(t)}{t} \quad \text{as } t \rightarrow \infty \quad (\text{A.1.1})$$

$$\begin{aligned}
1 - F(t) &= p \left(\frac{\sigma_g}{t - \mu_g} \phi\left(\frac{t - \mu_g}{\sigma_g}\right) \right) + q \left(\frac{\sigma_b}{t - \mu_b} \phi\left(\frac{t - \mu_b}{\sigma_b}\right) \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{p}{t - \mu_g} e^{-\frac{(t - \mu_g)^2}{2\sigma_g^2}} + \frac{q}{t - \mu_b} e^{-\frac{(t - \mu_b)^2}{2\sigma_b^2}} \right) \\
&\stackrel{(a)}{=} \frac{1}{\sqrt{2\pi}} \frac{p}{t - \mu_g} e^{-\frac{(t - \mu_g)^2}{2\sigma_g^2}} (1 + o(1)) \quad \text{as } t \rightarrow \infty
\end{aligned}$$

where (a) is true since

$$\lim_{t \rightarrow \infty} \frac{\frac{\frac{q}{t - \mu_b} e^{-\frac{(t - \mu_b)^2}{2\sigma_b^2}}}{\frac{p}{t - \mu_g} e^{-\frac{(t - \mu_g)^2}{2\sigma_g^2}}} - 1}{1} = \lim_{t \rightarrow \infty} \frac{e^{-\frac{(t - \mu_b)^2}{2\sigma_b^2}}}{e^{-\frac{(t - \mu_g)^2}{2\sigma_g^2}}} = 0$$

assuming $\sigma_g > \sigma_b$. So taking in consideration condition (5.1.3):

$$\begin{aligned}
\frac{1 - F(t + xg(t))}{1 - F(t)} &= \frac{\frac{1}{\sqrt{2\pi}} \frac{p}{t + xg(t) - \mu_g} e^{-\frac{(t + xg(t) - \mu_g)^2}{2\sigma_g^2}} (1 + o(1))}{\frac{1}{\sqrt{2\pi}} \frac{p}{t - \mu_g} e^{-\frac{(t - \mu_g)^2}{2\sigma_g^2}} (1 + o(1))} \\
&= \frac{t - \mu_g}{t + xg(t) - \mu_g} e^{\frac{-(t + xg(t) - \mu_g)^2 + (t - \mu_g)^2}{2\sigma_g^2}} (1 + o(1)) \\
&= \frac{1}{1 + \frac{xg(t)}{t - \mu_g}} e^{-\frac{g(t)x(t - \mu_g)}{\sigma_g^2}} e^{-\frac{g^2(t)x^2}{2\sigma_g^2}} (1 + o(1)) =
\end{aligned}$$

By choosing $g(t) = \frac{\sigma_g^2}{t - \mu_g}$ as the strictly positive function for $t \rightarrow \infty$ we get

$$= \frac{1}{1 + \frac{x\sigma_g^2}{(t - \mu_g)^2}} e^{-x} e^{-\frac{\sigma_g^2 x^2}{2(t - \mu_g)^2}} (1 + o(1)) = e^{-x} \quad \text{as } t \rightarrow \infty$$

That conclude that the distribution function $F(x)$ belongs to the domain of attraction of Type I. Similar analysis can be found in [82], where some examples for convergence of sequences of independent random variables with the same mixed distribution is investigated.

A.2 Normalizing constants

Derivation of a_K and b_K :

From Theorem 2 we know that u_K is a sequence of real numbers such that $K(1 - F(u_K)) \rightarrow \tau$ as $K \rightarrow \infty$, therefore in our case:

$$1 - pF_g(u_K) - qF_b(u_K) \rightarrow \frac{1}{K} e^{-x}, \quad K \rightarrow \infty$$

where $\tau = e^{-x}$. The same way as the previous proof using (A.1.1) we obtain that

$$\begin{aligned}
&\left(\frac{p\sigma_g}{u_K - \mu_g} \phi\left(\frac{u_K - \mu_g}{\sigma_g}\right) \right) + \left(\frac{q\sigma_b}{u_K - \mu_b} \phi\left(\frac{u_K - \mu_b}{\sigma_b}\right) \right) \rightarrow \frac{1}{K} e^{-x} \\
&\frac{1}{\sqrt{2\pi}} \frac{p}{u_K - \mu_g} e^{-\frac{(u_K - \mu_g)^2}{2\sigma_g^2}} (1 + o(1)) \rightarrow \frac{1}{K} e^{-x}
\end{aligned}$$

the last step is true since $u_K \rightarrow \infty$ as $K \rightarrow \infty$, similar to the pervious proof.

$$-\frac{1}{2} \log 2\pi + \log p - \log(u_K - \mu_g) - \frac{(u_K - \mu_g)^2}{2\sigma_g^2} + \log K + x + o(1) \rightarrow 0 \quad (\text{A.2.1})$$

It follows at once that $(t - \mu_g)^2 / 2 \log K \rightarrow 1$, and hence

$$\log(u_K - \mu_g) = \frac{1}{2}(\log 2 + \log \log K) + o(1)$$

Putting this in (A.2.1), we obtain

$$\frac{(u_K - \mu_g)^2}{2\sigma_g^2} = -\frac{1}{2} \log 2\pi + \log p - \frac{1}{2}(\log 2 + \log \log K) + \log K + x + o(1)$$

or

$$\frac{(u_K - \mu_g)^2}{\sigma_g^2} = 2 \log K \left(1 + \frac{x - \frac{1}{2} \log \frac{4\pi}{p^2} - \frac{1}{2} \log \log K}{\log K} + o\left(\frac{1}{\log K}\right) \right)$$

and hence

$$\frac{(u_K - \mu_g)}{\sigma_g} = \sqrt{2 \log K} \left(1 + \frac{x - \frac{1}{2} \log \frac{4\pi}{p^2} - \frac{1}{2} \log \log K}{2 \log K} + o\left(\frac{1}{\log K}\right) \right)$$

so by using expansion we have,

$$u_K = \sigma_g \sqrt{2 \log K} \left(1 + \frac{x - \frac{1}{2} \log \frac{4\pi}{p^2} - \frac{1}{2} \log \log K}{2 \log K} + o\left(\frac{1}{\log K}\right) \right) + \mu_g$$

since we know that $u_K = x/a_K + b_K$ we conclude that

$$\begin{aligned} a_K &= \frac{\sqrt{2 \log K}}{\sigma_g} \\ b_K &= \sigma_g \left((2 \log K)^{1/2} - \frac{\log \log K + \log \frac{4\pi}{p^2}}{2(2 \log K)^{1/2}} \right) + \mu_g \end{aligned}$$

Appendix B

Capacity - Direct expression analysis

B.1 Capacity - analyzing change in group sizes

The system can be modeled in a different way, still taking in consideration our channel model, we can analyse the channel capacity by addressing separately to the different groups which are differ by their channel state and evaluate the maximum for each group separately. Therefore the system is modeled as a first order time-homogeneous Markov chain with $K + 1$ states, each state represent The number of users which are in G state or in B state. In other words, The state $S_i(n)$ represent the system state in discrete time n , where i users are in Bad channel state and $K - i$ users are in Good channel state.

The Markov chain is irreducible and aperiodic chain. We can notice that it is possible to go with positive probability from any state of the Markov chain to any other state in a finite number of steps since the chain is a complete graph. And it is easy to see that since each state has positive probability to remain at the same state all the states are aperiodic.

we will give now some properties for the system behavior which will be used later.

B.1.1 Transition probability matrix

The transition probability matrix P is given:

$$P_{ij} = \begin{cases} \sum_{n=0}^{\min\{i, K-i, j, K-j\}} \binom{K-i}{j-i+n} \binom{i}{n} \alpha^{j-i+n} (1-\alpha)^{K-j-n} \beta^n (1-\beta)^{i-n} & \text{if } i \leq j \\ \sum_{n=0}^{\min\{i, K-i, j, K-j\}} \binom{K-i}{n} \binom{i}{i-j+n} \alpha^n (1-\alpha)^{K-i-n} \beta^{i-j+n} (1-\beta)^{j-n} & \text{if } i > j \end{cases} \quad \text{for } 0 \leq i, j \leq K \quad (\text{B.1.1})$$

where α and β represent the probability that user changes his channel state from G to B or the opposite as shown in Figure 4.1.

The entries in the transition matrix are binomials probabilities, the probability to get from state i to state j depends on the number of users which in a given time n changed their channel state. The probability P_{ij} sums up all the probabilities of the cases that eventually lead to having j users in bad state and $K - j$ users in good state. Each case describes the number of user transition from good to bad state with success probability α and the number of user transition from bad to good state with success probability β in a given time n .

Proposition 1. *The number of transition cases possible, which begins in state i and ends in state j in a given time n , are $\min\{i, K - i, j, K - j\} + 1$.*

Proof. The initial state i represents that there are i users in the bad group and $K - i$ users in the good group and the target state j represent that there are j users in the bad group and $K - j$ users in the good group. For example, the number of possible transitions, for $K = 6, i = 2, j = 3$, is 3; the first one is the transition of one user from the good group to the bad group and none from the bad group to the good group, the second is the transition of two users from the good group to the bad group and one user from the bad group to the good group and the third one is the transition of three users from the good group to the bad group and two users from the bad group to the good group. Another example, if i or j are 0 or K we would have only one possible transition.

The expression $\min\{i, K - i\}$ defines the maximum transitions possible from the group regardless the target state j , there can't be more transition cases than the size of the smallest group. The expression $\min\{j, K - j\}$ defines the maximum transitions possible in order to obtain the new sizes of the groups, since again we cannot count more transition cases than the size of the smallest group. Finally the minimal value between the two should be chosen, so it would hold for both conditions. Hence, we get : $\min\{\min\{i, K - i\}, \min\{j, K - j\}\} = \min\{i, K - i, j, K - j\}$ \square

Let us consider the assumption that the transition probabilities in the good-bad channel are symmetric (α equals β).

The transition matrix P_{ij} under the condition $\alpha = \beta$, (B.1.1), can be written as follow:

$$P_{ij} = \begin{cases} \sum_{n=0}^{\min\{i, K-i, j, K-j\}} \binom{K-i}{j-i+n} \binom{i}{n} \alpha^{j-i+2n} (1-\alpha)^{K-j+i-2n} & \text{if } i \leq j \\ \sum_{n=0}^{\min\{i, K-i, j, K-j\}} \binom{K-i}{n} \binom{i}{i-j+n} \alpha^{i-j+2n} (1-\alpha)^{K-i+j-2n} & \text{if } i > j \end{cases} \quad \text{for } 0 \leq i, j \leq K$$

Looking on the structure of the equation above we can notice that $P_{ij} = P_{(K-i)(K-j)}$ assuming $i \leq j$ and without loss of generality. In both entries we get the probability:

$$\sum_{n=0}^{\min\{i, K-i, j, K-j\}} \binom{K-i}{j-i+n} \binom{i}{n} \alpha^{j-i+2n} (1-\alpha)^{K-j+i-2n}$$

This applies also to the case where $i > j$.

So the matrix has a structure as shown below:

$$P_{i,j} = \begin{pmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,K-1} & a_{0,K} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,K-1} & a_{1,K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1,K} & a_{1,K-1} & \cdots & a_{1,1} & a_{1,0} \\ a_{0,K} & a_{0,K-1} & \cdots & a_{0,1} & a_{0,0} \end{pmatrix}$$

Proposition 2. *The transition matrix P is a Centrosymmetric (Cross-Symmetric) Matrix.*

Proof. An $(n \times n)$ matrix $A = [A_{i,j}]$ is centrosymmetric when its entries satisfy $A_{i,j} = A_{n-i+1, n-j+1}$ for $1 \leq i, j \leq n$ ([83]). In our case, $0 \leq i, j \leq n$ and the matrix is $((n+1) \times (n+1))$, so we would get that the entries satisfy $A_{i,j} = A_{n-i, n-j}$ for $0 \leq i, j \leq n$. This consist with the structure of the transition matrix explained above. \square

Let us consider the $(n \times n)$ exchange matrix J , which satisfies $I^2 = J$ where I is the identity matrix, and it entries defined such that:

$$J_{i,j} = \begin{cases} 1 & \text{if } j = n - i \\ 0 & \text{if } j \neq n - i \end{cases} \quad \text{for } 0 \leq i, j \leq n$$

Proposition 3 ([83]). *An $(n \times n)$ matrix A is centrosymmetric if and only if $JA = AJ$.*

B.1.2 Stationary distribution

If a Markov chain is a time-homogeneous, so that the process is described by a single, time-independent matrix P_{ij} then the vector $\bar{\pi}$ is called a stationary distribution and it satisfies the global balance condition : $\bar{\pi} = \bar{\pi}P$.

In the case that the finite state Markov chain is irreducible and aperiodic, then the stationary distribution exist and it is unique [84].

The characteristics mentioned above exist in our system model. We will investigate our stationary distribution $\bar{\pi}$.

Theorem 9. *The stationary distribution $\bar{\pi}$ of the Markov chain with transition matrix P is a symmetric distribution around its middle.*

Proof. Let $\bar{\pi}$ be the unique stationary distribution of the Markov chain which hold the global balance condition:

$$\begin{aligned}\bar{\pi} &= \bar{\pi}P \\ \bar{\pi}J &\stackrel{(a)}{=} \bar{\pi}PJ \\ \bar{\pi}^* &\stackrel{(b)}{=} \bar{\pi}PJ \\ \bar{\pi}^* &\stackrel{(c)}{=} \bar{\pi}JP \\ \bar{\pi}^* &\stackrel{(d)}{=} \bar{\pi}^*P\end{aligned}$$

Where (a) since the equation hold we can to multiply both sides with the matrix J as defined earlier, (b) any vector \bar{v} multiplied by the matrix J turns to be the same vector \bar{v} in its revers order, which mean $v_i^* = v_{|v|-i}$, (c) P is a centrosymmetric matrix and according to proposition 3, $PJ = JP$ (d) same as (b).

As a result we can see that the vector $\bar{\pi}^*$ satisfies the global balance equation therefore in order not to contradict the singularity of the stationary distribution the obvious conclusion is that $\bar{\pi} = \bar{\pi}^*$. From this we can conclude that the stationary distribution is a symmetric distribution. \square

A random variable ξ and its distribution function $F(x)$ are said to be unimodal with mode M if $F(x)$ is convex on $(-\infty, M)$ and concave on (M, ∞) . The notion of unimodality has proved to be very fruitful, therefore its analogs have been introduced for discrete and multivariate distributions. In [85] some results for discrete unimodality can be found, we will consider some of these results for better understanding the stationary distribution $\bar{\pi}$. A distribution $\{p_n\}$ with $k \in \mathbb{Z}$ as its support will be said to be unimodal if there exists at least one integer M such that:

$$\begin{aligned}p_n &\geq p_{n-1}, \quad \text{for all } n \leq M, \\ p_{n+1} &\leq p_n, \quad \text{for all } n \geq M.\end{aligned}$$

The distribution may consist of a single point mass, or it may be bounded on one side, both sides, or be unbounded. A discrete distribution $\{h_n\}$ is strongly unimodal if the convolution of $\{h_n\}$ with any unimodal $\{p_n\}$ is unimodal. In addition, a necessary and sufficient condition that $\{h_n\}$ is strongly unimodal is that ([85], *Theorem 3*):

$$h_n^2 \geq h_{n-1}h_{n+1} \quad \text{for all } n$$

Another theorem presented is: The limit of a convergent sequence of strongly unimodal discrete distributions is strongly unimodal ([85], *Theorem 1*).

Also, it has been proven that any binomial distribution is strongly unimodal since it is a convolution of a finite number of Bernoulli distributions which are strongly unimodal. This derives from the *Lemma* given in [85], that the convolution of two strongly unimodal discrete distributions is strongly unimodal.

Theorem 10. *The stationary distribution $\bar{\pi}$ of the Markov chain with transition matrix P is a discrete strongly unimodal distribution.*

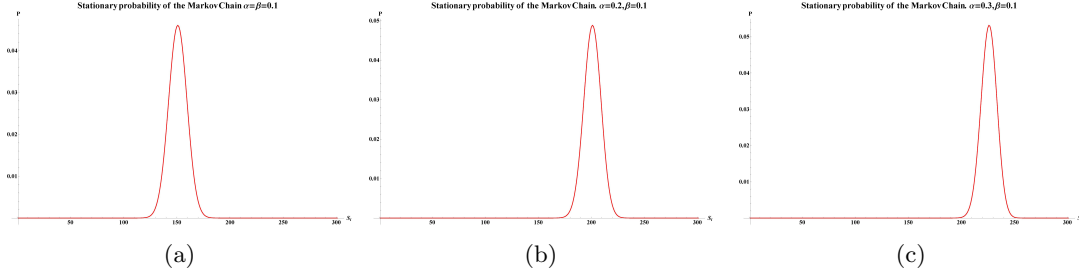


Figure B.1: Stationary distribution of the Markov chain system model. Where the transition probabilities in (a) $\alpha = \beta = 0.1$ (b) $\alpha = 0.2$ $\beta = 0.1$ (c) $\alpha = 0.3$ $\beta = 0.1$

Proof. Since the stationary distribution is defined as the limit distribution in an ergodic Markov chain ([84], *Theorem 59*), according to ([85], *Theorem 1*), in order to prove that $\bar{\pi}$ is a discrete strongly unimodal distribution it is sufficient to prove that each row of the transition matrix P is strongly unimodal.

Let us consider the following situation to describe the system model. Given the system state is i (i users in the bad group and $K - i$ users in the good group), we define two binomial random variables, $X \sim B(K - i, \alpha)$ and $Y \sim B(i, \alpha)$ which represent the number of users that traversed from the good group to the bad group and from the bad group to the good group in a time unit, respectively. Therefore, the distribution for the next system state J , i.e. the row i of the transition matrix P is:

$$Pr(J = j) = Pr(i + X - Y = j) = Pr(X - Y = j - i)$$

The next system state is ruled from the given state i the amount of users that traversed to the bad group minus the users that traversed to the good group. The difference between the two binomial random variables X and Y gives the next state distribution, hence we only need to show that this distribution is strongly unimodal. In the case of sum of two independent binomial random variables it is very simple, since each of them has strongly unimodal distribution and the distribution of the sum is the convolution of the two strongly unimodal distributions which is strongly unimodal. Let us define a new random variable W such that $W = -Y$. The distribution of W is defined:

$$Pr(W = -k) = \binom{i}{k} \alpha^k (1 - \alpha)^{i-k} \quad \text{for } k = 0, 1, 2, \dots, i$$

This distribution is clearly strongly unimodal since the necessary and sufficient condition in ([85], *Theorem 3*) holds. For Y , from its binomial characteristics, the inequality $p_k^2 \geq p_{k+1}p_{k-1}$ holds. Since $p(Y = k) = p(W = -k)$, $p(Y = k + 1) = p(W = -k - 1)$ and $p(Y = k - 1) = p(W = -k + 1)$, the inequality $p_{-k}^2 \geq p_{-k+1}p_{-k-1}$ for W holds as well.

Therefore, we get that the distribution:

$$Pr(J = j) = Pr(X - Y = j - i) = Pr(X + W = j - i)$$

is strongly unimodal, since X and W each has a strongly unimodal distribution, the convolution between them will also give a strongly unimodal distribution [85]. \square

From Theorems 9 and 10, we can conclude that the stationary distribution of the Markov chain is a symmetric unimodal function. See Figure B.1a for the case of symmetric transition probabilities. We will derive the capacity for this scenario given the current state of the system, while this state determines the amount of users experiencing the different channel distributions.

B.1.3 Capacity analyze

We consider the problem of finding $\widetilde{M_K}$ as defined in (5.0.1) and the distribution of $\{C_i\}$ as defined (4.1.1). As mentioned earlier at first we will assume that the system is in a specific state $S_i(n)$ and

therefore we know the amount of users which are in good and bad channel state. It is important to understand that given the system state $S_i(n)$, there are i users in Bad channel state which we denote by K_1 , and $K - i$ users in Good channel state which we denote by K_2 , note that $K = K_1 + K_2$. In order to find the expected capacity we shall derive the capacity distribution of \widetilde{M}_K :

$$\begin{aligned} P_r(\widetilde{M}_K < x) &= \sum_{i=0}^K P_r(\widetilde{M}_K < x \mid S_i(n)) P(S_i(n)) = \\ &= \sum_{i=0}^K P_r(\max\{C_1(n), C_2(n), \dots, C_K(n)\} < x \mid S_i(n)) P(S_i(n)) \end{aligned} \quad (\text{B.1.2})$$

Given $S_i(n)$ we get two independent groups $A = \{a_1, \dots, a_{K_1}\}$ and $B = \{b_1, \dots, b_{K_2}\}$ where a_j and b_j represents the users' capacities according to the group each user belongs to. Since every state of the system need to be taken in consideration, we cannot use the EVT for all the group sizes. Let us define a parameter φ for the minimal group size which we can use EVT to achieve the asymptotic distribution for the maximum capacity. Assuming A and B are big enough, the distribution of the maximum capacity in each group is the asymptotic distribution $G_j(x, K_i)$ as in equation (3.3.3), where the parameters a_K and b_K will be a function of K_1 and K_2 . And for the states where one group is too small for us to use EVT we will treat in a straightforward manner. The distribution given the system state $S_i(n)$:

$$\begin{aligned} P_r(\widetilde{M}_K < x \mid S_i(n)) &= P_r(\max\{C_1(n), C_2(n), \dots, C_K(n)\} < x \mid S_i(n)) \\ &= P_r(\max\{\max\{A\}, \max\{B\}\} < x \mid S_i(n)) \\ &= P_r(\max\{A\} < x, \max\{B\} < x \mid S_i(n)) \\ &\stackrel{(a)}{=} P_r(\max\{A\} < x \mid S_i(n)) P_r(\max\{B\} < x \mid S_i(n)) \end{aligned}$$

where (a) is due to the independency between the groups. So (B.1.2) can be written:

$$\begin{aligned} P_r(\widetilde{M}_K < x) &= \sum_{i=0}^K P_r(\max\{A\} < x \mid S_i(n)) P_r(\max\{B\} < x \mid S_i(n)) P(S_i(n)) \\ &= \sum_{i=0}^{\varphi} F_B^{K_1}(x) G_2(x, K_2) + \sum_{i=\varphi+1}^{K-\varphi} G_1(x, K_1) G_2(x, K_2) + \sum_{i=K-\varphi+1}^K G_1(x, K_1) F_G^{K_2}(x) \end{aligned} \quad (\text{B.1.3})$$

Thus the expected channel capacity is:

$$\begin{aligned} E[\widetilde{M}_K] &= \sum_{i=0}^K E[\widetilde{M}_K \mid S_i(n)] P(S_i(n)) \\ &= \sum_{i=0}^{\varphi} P(S_i(n)) \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} [F_B^{K_1}(x) G_2(x, K_2)] dx \\ &\quad + \sum_{i=\varphi+1}^{K-\varphi} P(S_i(n)) \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} [G_1(x, K_1) G_2(x, K_2)] dx \\ &\quad + \sum_{i=K-\varphi+1}^K P(S_i(n)) \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} [G_1(x, K_1) F_G^{K_2}(x)] dx \end{aligned} \quad (\text{B.1.4})$$

In Figure B.2 we can see the capacity PDF in the case $N_1 = 150$ and $N_2 = 150$ where each group experience different channel distribution with $\mu_1 = 1.25, \mu_2 = 0.25$ and $\sigma_1 = \sigma_2 = 0.19$ respectively.

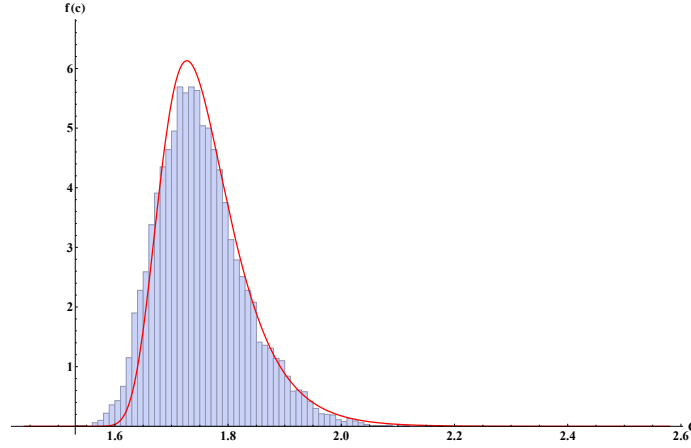
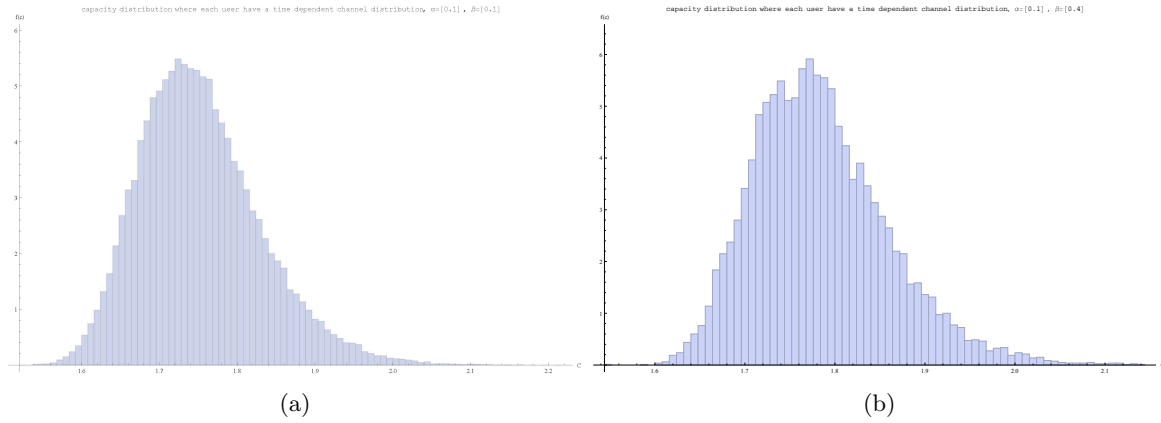


Figure B.2: Capacity Distribution with Two Groups distributed differently


 Figure B.3: Capacity distribution for variable size of Groups distributed differently. Where the transition probabilities in (a) $\alpha = \beta = 0.1$ (b) $\alpha = 0.1$ $\beta = 0.4$

A more realistic simulation is in Figure B.3, we can see the maximal MIMO capacity among 300 users, with time variable size for each group (bad or good), and time variable state for each user according to the good bad channel Markov model. Achieved by saving a state vector for the users and updating it following the transition probabilities for each drawn.

B.1.4 Capacity Bounds

Finding a closed solution for the expected channel capacity for the expression given in (B.1.4) is very complicated, because each $C_i(n)$ is determined according to a Markov process. One can think of some upper and lower bounds for the expected channel capacity in the analytical development in equation (B.1.4), the first naive lower bound would be to choose i such that gives the maximum probability from the stationary distribution $\bar{\pi}$ of the system model (figure B.1c) in order to simplify the expression in (B.1.4), once we know the specific i we know the parameters K_1 and K_2 :

$$\begin{aligned}
 E[\widetilde{M}_K] &\stackrel{(a)}{=} \sum_{i=0}^K E[\widetilde{M}_K \mid \pi_i] \pi_i \\
 &\stackrel{(b)}{\geq} \max_i \{\pi_i\} E[\widetilde{M}_K \mid \pi_i]
 \end{aligned} \tag{B.1.5}$$

where (a) is from equation (B.1.4) and (b) is to choose the maximum probability in $\bar{\pi}$, note that $P(S_i(n)) = \pi_i$.

The index i which satisfy the $\max_i \{\pi_i\}$ in the case where the transition probabilities in the good-bad channel are symmetric is $i = \frac{K}{2}$ as shown earlier in the analysis of the stationary distribution of the system.

Since $i = \frac{K}{2}$ which mean that the groups are still big as $K \rightarrow \infty$, its fair to say that the EVT can be applied on the distribution of the two groups.

Therefore, we can achieve a closed bound for this case by setting $i = \frac{K}{2}$ in equation (B.1.5):

$$\begin{aligned}
 E[\widetilde{M}_K] &\geq \pi_{\frac{K}{2}} E[\widetilde{M}_K \mid \pi_{\frac{K}{2}}] \\
 &= \pi_{\frac{K}{2}} \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} \left[G_1(x, \frac{K}{2}) G_2(x, \frac{K}{2}) \right] dx \\
 &\stackrel{(a)}{=} \pi_{\frac{K}{2}} \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} \left[(G_1(x, \frac{K}{2}))^2 \right] dx \\
 &= \pi_{\frac{K}{2}} \int_{-\infty}^{\infty} x \frac{\partial}{\partial x} \left[\left(e^{-e^{-\frac{x-b\frac{K}{2}}}{a\frac{K}{2}}} \right)^2 \right] dx \\
 &= 2\pi_{\frac{K}{2}} \int_{-\infty}^{\infty} x e^{-e^{-\frac{x-b\frac{K}{2}}}{a\frac{K}{2}}} e^{-\frac{x-b\frac{K}{2}}{a\frac{K}{2}}} \frac{1}{a\frac{K}{2}} dx \\
 &\stackrel{(b)}{=} 2\pi_{\frac{K}{2}} \int_0^{\infty} (a_{\frac{K}{2}} \ln(\frac{1}{y}) + b_{\frac{K}{2}}) e^{-y} dy \\
 &= 2\pi_{\frac{K}{2}} (a_{\frac{K}{2}} \gamma + b_{\frac{K}{2}})
 \end{aligned} \tag{B.1.6}$$

where $a_{\frac{K}{2}}$ and $b_{\frac{K}{2}}$ as defined in (3.3.6) and (3.3.7).

The first lower bound given above can be improved. In this bound we only took in consideration the maximum probability from the sum of probabilities in equation (B.1.3). As mentioned we know that the two groups sizes are $O(K/2)$, hence EVT can be used, and another bound can be taken in consideration which includes only the states that represent similar group sizes. Let us consider the situation that there is more good users than bad users, i.e. $\frac{K}{2} - \delta$ bad users and $\frac{K}{2} + \delta$ good users, where δ can be chosen as we please in order to maximize the offered bound. Now if we discard 2δ users from the good group we would get a lower bound for the expected channel capacity, since the users who give the most contribution are the good users, so it is clear that if we reduce these users

we reduce the expected channel capacity.

$$\begin{aligned}
 E[\widetilde{M}_K] &= \sum_{i=0}^K E[\widetilde{M}_K \mid \pi_i] \pi_i \\
 &\geq \sum_{i=\frac{K}{2}-\delta}^{\frac{K}{2}} E[\widetilde{M}_K \mid \pi_i] \pi_i \\
 &= \sum_{i=\frac{K}{2}-\delta}^{\frac{K}{2}} E[G_1(x, K_1)G_2(x, K_2)] \pi_i \\
 &\geq \sum_{i=\frac{K}{2}-\delta}^{\frac{K}{2}} E[G_1(x, K_1)G_2(x, K_1)] \pi_i \\
 &= \sum_{i=\frac{K}{2}-\delta}^{\frac{K}{2}} (a_{K_1}\gamma + b_{K_1}) \pi_i \\
 &\geq \left(a_{\frac{K}{2}-\delta}\gamma + b_{\frac{K}{2}-\delta}\right) \pi_{\frac{K}{2}-\delta}\delta
 \end{aligned} \tag{B.1.7}$$

Appendix C

Random selection Proof

Let N be a Poisson process with intensity λ and N^* be a point process obtained from N such that each event is removed or retained independently with probabilities $1 - p$ and p respectively. So for any Borel set B ,

$$\begin{aligned} P(N^*(B) = k) &= \sum_{i=k}^{\infty} P(N(B) = i) P(N^*(B) = k | N(B) = i) \\ &= \sum_{i=k}^{\infty} \frac{e^{-\lambda(B)} (\lambda(B))^i}{i!} \binom{i}{k} p^k (1-p)^{i-k} \\ &= \frac{e^{-p\lambda(B)} (p\lambda(B))^k}{k!} \end{aligned}$$

so that $N^*(B)$ is a Poisson random variable with mean $p\lambda(B)$.

Appendix D

System-status chain transition probabilities

Here we show the calculation of the transition probabilities of the system chain. We use the guide lines in [3] to simplify things and use also their notions but under our case which state that $r_i = s_i = q_i = p_i$. Due to this fact there are many more possible transition in the state space and therefore the following calculations would be more complicated. In the same manner the calculations are based on the change in the number of active users, thus the transitions can be classified into four types as follows:

1. The number of active users transits from 0 to 1:

Only a blocked user may become active, and only one, say j . All other blocked users must not exceed the threshold. In addition all idle users which may receive a package must not exceed the threshold and therefore, say w , becomes blocked. The latter will repeat it self during this calculation. The probability of such a transition is expressed by:

$$P(\Delta A = 1, \Delta B = -1+w, \Delta I = -w) = \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \frac{p_j}{\bar{p}_j} [(1 - P_j(0|2)) + \lambda_j P_j(0|2)]$$

where the product is on the sets of users by their transition and n is the number of blocked users after the transition. Note that the user j will be active if his queue is not empty after the successful transmission or a package arrived in the beginning of the slot. Also note that \bar{p} and $\bar{\lambda}$ stand for $1 - p$ and $1 - \lambda$, respectively.

2. The number of active users remains 1:

This case divides to two subcases, the same active user j remains active or another blocked user s becomes active. Nevertheless w idle users may still become blocked.

- (a) the probability for the first case is:

$$P(\Delta A = 0, \Delta B = +w, \Delta I = -w, j \rightarrow j) = \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \frac{p_j}{\bar{p}_j} [P_j(1|1) + \lambda_j(1 - P_j(1|1))]$$

- (b) the probability for the second case is:

$$P(\Delta A = 0, \Delta B = +w, \Delta I = -w, j \rightarrow s) = \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \frac{p_s}{\bar{p}_s} [P_s(1|1) + \lambda_s(1 - P_s(1|1))]$$

3. The number of active users transits from 1 to 0:

This case is subdivided to three subcases,

- (a) The first subcase is that the active user j becomes idle:

$$P(\Delta A = -1, \Delta B = +w, \Delta I = -w) = \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \frac{p_j}{\bar{p}_j} [\bar{\lambda}_j (1 - P_j(1|1))]$$

- (b) The second subcase is that the active user j becomes blocked and blocked user s becomes idle:

$$P(\Delta A = -1, \Delta B = +1-1+w, \Delta I = +1-w) = \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \frac{p_s}{\bar{p}_s} [\bar{\lambda}_s (1 - P_s(1|1))]$$

- (c) The third subcase is that the active user j becomes blocked:

$$\begin{aligned} P(\Delta A = -1, \Delta B = +1 + w, \Delta I = -w) = & \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \\ & + \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \sum_s^{K-n-w-1} \frac{\lambda_s p_s}{\bar{\lambda}_s} \prod_i^{n+1} \bar{p}_i \\ & + \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \left(1 - \prod_s^w \bar{p}_s\right) \left(1 - \prod_i^n \bar{p}_i\right) \\ & + \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \left[1 - \prod_s^n \bar{p}_s - \bar{p}_j \prod_i^n \bar{p}_i \sum_q^n \frac{p_q}{\bar{p}_q}\right] \\ & + \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_i^n \bar{p}_i \prod_l^w \lambda_l \left[1 - \prod_s^w \bar{p}_s - \bar{p}_j \prod_b^w \bar{p}_b \sum_q^w \frac{p_q}{\bar{p}_q}\right] \end{aligned}$$

The first expression is for the case which non of the users exceeds the threshold. The second is for the case which one idle user managed to successfully transmits. The third case describes the case which one or more from the blocked and the idle groups tries to transmit, therefore what happens with user j is meaningless. The forth expression describes the situation which user j collided with one or more users from the blocked group, or collision happened between the blocked users and j didn't exceed the threshold. The fifth expression is the same as the fourth concerning the group of idle users which receives a package.

4. The number of active users remains 0 :

This case is subdivided to three subcases,

- (a) The first subcase is that all users maintain their status without change:

$$P(\Delta B = 0, \Delta I = 0) = \left[1 - \prod_i^n \bar{p}_i \sum_q^n \frac{p_q}{\bar{p}_q}\right] \prod_k^{K-n} \bar{\lambda}_k + \prod_i^n \bar{p}_i \prod_k^{K-n} \bar{\lambda}_k \sum_q^{K-n} \frac{\lambda_q p_q}{\bar{\lambda}_q}$$

where the first term is the probability that no packet arrives at the idle users while no blocked user, or at least two blocked users, transmit, and the second term is the probability that no blocked user exceeds the threshold while only one of the idle users successfully transmits.

- (b) The second subcase is that one blocked user j becomes idle:

$$P(\Delta B = -1 + w, \Delta I = +1 - w) = \prod_k^{K-n-w-1} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \prod_i^{n+1} \bar{p}_i \left[\bar{\lambda}_j \frac{p_j}{\bar{p}_j} P_j(0|2)\right]$$

(c) The third subcase describes some situation for w idle users becomes blocked:

$$\begin{aligned}
P(\Delta B = +w, \Delta I = -w) = & \\
& \prod_k^{K-n-w} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \left[1 - \prod_i^n \bar{p}_i \sum_q^n \frac{p_q}{\bar{p}_q} \right] \\
& + \prod_k^{K-n-w} \bar{\lambda}_k \prod_l^w \lambda_l \left[1 - \prod_s^w \bar{p}_s - \prod_b^w \bar{p}_b \sum_q^w \frac{p_q}{\bar{p}_q} \right] \\
& + \prod_k^{K-n-w} \bar{\lambda}_k \prod_l^w \lambda_l \bar{p}_l \sum_q^w \frac{p_q}{\bar{p}_q} \left[1 - \prod_i^n \bar{p}_i \right] \\
& + \prod_k^{K-n-w} \bar{\lambda}_k \sum_s^{K-n-w} \frac{\lambda_s p_s}{\bar{\lambda}_s} \prod_l^w \lambda_l \bar{p}_l \prod_i^n \bar{p}_i
\end{aligned}$$

The first expression is for the case which non of the w users exceeds the threshold and no blocked user, or at least two blocked users, transmit, the second is for the case which at least two users from w exceeds the threshold, the third case describes collision between one of the users from w with at least one from the blocked users and the forth expression describes the situation which one idle user succussed to transmit while all the other doesn't exceed the threshold.

According to these transition probabilities we can calculate the steady state of the chain given the auxiliary quantities $p_i(1|1)$ and $p_i(0|2)$, the probability of exceedance and the users arrival rates.

Appendix E

Average success probabilities

Once the steady state of the system chain is known, the average success probabilities can be calculated. We do so by calculate the values of $P_B(i)$, $P_A(i)$ and $P_I(i)$ in the same manner as calculated in [3] while taking in consideration that user must exceed the threshold in order to transmit.

1. For Blocked user i :

$$P_B(i) = Pr(\text{user } i \text{ success} \mid \text{user } i \text{ is blocked}) = \frac{Pr(\text{user } i \text{ success and user } i \text{ is blocked})}{Pr(\text{user } i \text{ is blocked})}$$

where

$$Pr(\text{user } i \text{ success and user } i \text{ is blocked}) = p_i \sum_{\substack{S_i=2 \\ j \neq i \\ S_j=0,1,2}} P(S_1, \dots, S_K) \prod_{j \neq i} (\lambda_j \bar{p}_j + \bar{\lambda}_j)^{\delta_{S_j=0}} (\bar{p}_j)^{\delta_{S_j=1}} (\bar{p}_j)^{\delta_{S_j=2}}$$

The exponent δ equals 1 when it's condition holds.

$$Pr(\text{user } i \text{ is blocked}) = \sum_{S_i=2} P(S_1, \dots, S_i, \dots, S_K)$$

2. For Active user i :

$$P_A(i) = Pr(\text{user } i \text{ success} \mid \text{user } i \text{ is active}) = \frac{Pr(\text{user } i \text{ success and user } i \text{ is active})}{Pr(\text{user } i \text{ is active})}$$

where

$$Pr(\text{user } i \text{ success and user } i \text{ is active}) = p_i \sum_{\substack{S_i=1 \\ j \neq i \\ S_j=0,2}} P(S_1, \dots, S_K) \prod_{j \neq i} (\lambda_j \bar{p}_j + \bar{\lambda}_j)^{\delta_{S_j=0}} (\bar{p}_j)^{\delta_{S_j=2}}$$

the difference here is due to the fact that no more than one user may be active.

$$Pr(\text{user } i \text{ is active}) = \sum_{S_i=1} P(S_1, \dots, S_i, \dots, S_K)$$

3. For Idle user i :

$$P_I(i) = Pr(\text{user } i \text{ success} \mid \text{user } i \text{ is idle}) = \frac{Pr(\text{user } i \text{ success and user } i \text{ is idle})}{Pr(\text{user } i \text{ is idle})}$$

where

$$Pr(\text{user } i \text{ success and user } i \text{ is idle}) = \lambda_i p_i \sum_{\substack{S_i=0 \\ j \neq i \\ S_j=0,1,2}} P(S_1, \dots, S_K) \prod_{j \neq i} (\lambda_j \bar{p}_j + \bar{\lambda}_j)^{\delta_{S_j=0}} (\bar{p}_j)^{\delta_{S_j=1}} (\bar{p}_j)^{\delta_{S_j=2}}$$

and

$$Pr(\text{user } i \text{ is idle}) = \sum_{S_i=0} P(S_1, \dots, S_i, \dots, S_K)$$

Using the above success probability it is not difficult to attain the expression of (6.1.6), (6.1.7), (6.1.8), (6.1.9), (6.1.10) from the steady state of the queue-length Markov chain. After knowing these values the boundary conditions can be calculated:

$$P_i(1 | 1) = 1 - \frac{\pi(1, 1)}{G_1^i(1) - \pi(1, 0)}$$

$$P_i(0 | 2) = \frac{\pi(0, 0)}{G_0^i(1)}$$

Bibliography

- [1] Jinsu Kim, Sungwoo Park, Jae Hong Lee, Joonho Lee, and Hanwook Jung. A scheduling algorithm combined with zero-forcing beamforming for a multiuser mimo wireless system. In *IEEE Vehicular Technology Conference*, volume 1, pages 211–215, 2005.
- [2] Taesang Yoo and Andrea Goldsmith. On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. *Selected Areas in Communications, IEEE Journal on*, 24(3):528–541, 2006.
- [3] Anthony Ephremides and Rong-Zhu Zhu. Delay analysis of interacting queues with an approximate model. *Communications, IEEE Transactions on*, 35(2):194–201, 1987.
- [4] Emre Telatar. Capacity of multi-antenna gaussian channels. *European transactions on telecommunications*, 10(6):585–595, 1999.
- [5] Andrea Goldsmith, Syed Ali Jafar, Nihar Jindal, and Sriram Vishwanath. Capacity limits of mimo channels. *Selected Areas in Communications, IEEE Journal on*, 21(5):684–702, 2003.
- [6] Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005.
- [7] Gerard J Foschini and Michael J Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless personal communications*, 6(3):311–335, 1998.
- [8] Raymond Knopp and Pierre A Humblet. Information capacity and power control in single-cell multiuser communications. In *Communications, 1995. ICC'95 Seattle, 'Gateway to Globalization', 1995 IEEE International Conference on*, volume 1, pages 331–335. IEEE, 1995.
- [9] Ido Bettesh and Shlomo Shamai. A low delay algorithm for the multiple access channel with rayleigh fading. In *Personal, Indoor and Mobile Radio Communications, 1998. The Ninth IEEE International Symposium on*, volume 3, pages 1367–1372. IEEE, 1998.
- [10] Jinho Choi and Fumiyuki Adachi. User selection criteria for multiuser systems with optimal and suboptimal lr based detectors. *Signal Processing, IEEE Transactions on*, 58(10):5463–5468, 2010.
- [11] Hanan Weingarten, Yossef Steinberg, and Shlomo Shamai. The capacity region of the gaussian multiple-input multiple-output broadcast channel. *Information Theory, IEEE Transactions on*, 52(9):3936–3964, 2006.
- [12] Masoud Sharif and Babak Hassibi. A comparison of time-sharing, dpc, and beamforming for mimo broadcast channels with many users. *Communications, IEEE Transactions on*, 55(1):11–15, 2007.
- [13] Randa Zakhour and Stephen Hanly. Min-max fair coordinated beamforming via large systems analysis. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 1990–1994. IEEE, 2011.

- [14] Pramod Viswanath, David N. C. Tse, and Rajiv Laroia. Opportunistic beamforming using dumb antennas. *Information Theory, IEEE Transactions on*, 48(6):1277–1294, 2002.
- [15] Giuseppe Caire. Mimo downlink joint processing and scheduling: a survey of classical and recent results. In *Proc. Workshop on Information Theory and Its Applications*. Citeseer, 2006.
- [16] Man-On Pun, Visa Koivunen, and H Vincent Poor. Opportunistic scheduling and beamforming for mimo-sdma downlink systems with linear combining. In *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, pages 1–6. IEEE, 2007.
- [17] Chiung-Jang Chen and Li-Chun Wang. Enhancing coverage and capacity for multiuser mimo systems by utilizing scheduling. *Wireless Communications, IEEE Transactions on*, 5(5):1148–1157, 2006.
- [18] Manish Airy, S Shakkattai, and Robert W Heath Jr. Spatially greedy scheduling in multi-user mimo wireless systems. In *Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on*, volume 1, pages 982–986. IEEE, 2003.
- [19] Young-June Choi, Jongtaek Kim, and Saewoong Bahk. Downlink scheduling with fairness and optimal antenna assignment for mimo cellular systems. In *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, volume 5, pages 3165–3169. IEEE, 2004.
- [20] Giuseppe Primolevo, Osvaldo Simeone, and Umberto Spagnolini. Channel aware scheduling for broadcast mimo systems with orthogonal linear precoding and fairness constraints. In *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, volume 4, pages 2749–2753. IEEE, 2005.
- [21] Taesang Yoo, Nihar Jindal, and Andrea Goldsmith. Finite-rate feedback mimo broadcast channels with a large number of users. In *Information Theory, 2006 IEEE International Symposium on*, pages 1214–1218. IEEE, 2006.
- [22] Babak Hassibi and Masoud Sharif. Fundamental limits in mimo broadcast channels. *Selected Areas in Communications, IEEE Journal on*, 25(7):1333–1344, 2007.
- [23] Guocong Song, Ye Li, Leonard J Cimini Jr, and Haitao Zheng. Joint channel-aware and queue-aware data scheduling in multiple shared wireless channels. In *Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE*, volume 3, pages 1939–1944. IEEE, 2004.
- [24] Atilla Eryilmaz and R Srikant. Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 3, pages 1794–1803. IEEE, 2005.
- [25] Charles Swannack, Elif Uysal-Biyikoglu, and Gregory Wornell. Low complexity multiuser scheduling for maximizing throughput in the mimo broadcast channel. In *Proc. Allerton Conf. Communications, Control and Computing*, 2004.
- [26] Li-Chun Wang, Cheng-Wei Chiu, Chu-Jung Yeh, and Chi-Fang Li. Coverage enhancement for ofdm-based spatial multiplexing systems by scheduling. In *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, pages 1439–1443. IEEE, 2007.
- [27] Yonghe Liu and Edward Knightly. Opportunistic fair scheduling over multiple wireless channels. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 1106–1115. IEEE, 2003.
- [28] Yangyang Zhang, Chunlin Ji, Yi Liu, W Malik, D O'Brien, and D Edwards. A low complexity user scheduling algorithm for uplink multiuser mimo systems. *Wireless Communications, IEEE Transactions on*, 7(7):2486–2491, 2008.

- [29] Bernd Bandemer and Samuli Visuri. Capacity-based uplink scheduling using long-term channel knowledge. In *Communications, 2007. ICC'07. IEEE International Conference on*, pages 785–790. IEEE, 2007.
- [30] Byong Lee, Hui Je, Oh-Soon Shin, and Kwang Lee. A novel uplink mimo transmission scheme in a multicell environment. *Wireless Communications, IEEE Transactions on*, 8(10):4981–4987, 2009.
- [31] Xiangping Qin and Randall Berry. Exploiting multiuser diversity for medium access control in wireless networks. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 1084–1094. IEEE, 2003.
- [32] Xiangping Qin and Randall A Berry. Distributed approaches for exploiting multiuser diversity in wireless networks. *Information Theory, IEEE Transactions on*, 52(2):392–413, 2006.
- [33] Kai Bai and Junshan Zhang. Opportunistic multichannel aloha: distributed multiaccess control scheme for ofdma wireless networks. *Vehicular Technology, IEEE Transactions on*, 55(3):848–855, 2006.
- [34] Xiangping Qin and Randall A Berry. Distributed power allocation and scheduling for parallel channel wireless networks. *Wireless Networks*, 14(5):601–613, 2008.
- [35] Xiangping Qin and Randall Berry. Opportunistic splitting algorithms for wireless networks. In *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 3, pages 1662–1672. IEEE, 2004.
- [36] Toan To and Jinho Choi. On exploiting idle channels in opportunistic multichannel aloha. *Communications Letters, IEEE*, 14(1):51–53, 2010.
- [37] Andres Garcia-Saavedra, Albert Banchs, Pablo Serrano, and Joerg Widmer. Distributed opportunistic scheduling: A control theoretic approach. In *INFOCOM, 2012 Proceedings IEEE*, pages 540–548. IEEE, 2012.
- [38] Joseph Kampeas, Asaf Cohen, and Omer Gurewitz. Capacity of distributed opportunistic scheduling in heterogeneous networks. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pages 1246–1253. IEEE, 2012.
- [39] Wan Choi and Jeffrey G Andrews. The capacity gain from intercell scheduling in multi-antenna systems. *Wireless Communications, IEEE Transactions on*, 7(2):714–725, 2008.
- [40] Weiyao Xiao and David Starobinski. Extreme value fec for reliable broadcasting in wireless networks. *Selected Areas in Communications, IEEE Journal on*, 28(7):1180–1189, 2010.
- [41] Weiyao Xiao, Sachin Agarwal, David Starobinski, and Ari Trachtenberg. Reliable rateless wireless broadcasting with near-zero feedback. *Networking, IEEE/ACM Transactions on*, 20(6):1924–1937, 2012.
- [42] Edgar N Gilbert et al. Capacity of a burst-noise channel. *Bell Syst. Tech. J*, 39(9):1253–1265, 1960.
- [43] EO Elliott. Estimates of error rates for codes on burst-noise channels. *Bell Syst. Tech. J*, 42(9):1977–1997, 1963.
- [44] Qinqing Zhang and Saleem A Kassam. Finite-state markov model for rayleigh fading channels. *Communications, IEEE Transactions on*, 47(11):1688–1692, 1999.
- [45] Ezio Biglieri, John Proakis, and Shlomo Shamai. Fading channels: Information-theoretic and communications aspects. *Information Theory, IEEE Transactions on*, 44(6):2619–2692, 1998.

- [46] David N. C. Tse and Stephen V Hanly. Multiaccess fading channels. i. polymatroid structure, optimal resource allocation and throughput capacities. *Information Theory, IEEE Transactions on*, 44(7):2796–2815, 1998.
- [47] Stephen V Hanly and David N. C. Tse. Multiaccess fading channels. ii. delay-limited capacities. *Information Theory, IEEE Transactions on*, 44(7):2816–2831, 1998.
- [48] Boris Gnedenko. Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943.
- [49] Jacques Janssen. Les processus (jx). *Cahiers du CERO*, 11(1969):181–214, 1969.
- [50] Sidney I Resnick and Marcel F Neuts. Limit laws for maxima of a sequence of random variables defined on a markov chain. *Advances in Applied Probability*, pages 323–343, 1970.
- [51] Sidney I Resnick and RJ Tomkins. Almost sure stability of maxima. *Journal of Applied Probability*, pages 387–401, 1973.
- [52] Sidney I Resnick. Stability of maxima of random variables defined on a markov chain. *Advances in Applied Probability*, pages 285–295, 1972.
- [53] GE Denzel and GL O’Brien. Limit theorems for extreme values of chain-dependent processes. *The Annals of Probability*, pages 773–779, 1975.
- [54] KF Turkman and MF Oliveira. Limit laws for the maxima of chain-dependent sequences with positive extremal index. *Journal of applied probability*, pages 222–227, 1992.
- [55] KF Turkman and AM Walker. Limit laws for the maxima of a class of quasi-stationary sequences. *Journal of applied probability*, pages 814–821, 1983.
- [56] MR Leadbetter. On extreme values in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 28(4):289–303, 1974.
- [57] William P McCormick and Lynne Seymour. Rates of convergence and approximations to the distribution of the maximum of chain-dependent sequences. *Extremes*, 4(1):23–52, 2001.
- [58] Guy Fayolle and Roudolf Iasnogorodski. Two coupled processors: the reduction to a riemann-hilbert problem. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 47(3):325–351, 1979.
- [59] Leonard Kleinrock and Yechiam Yemini. Interfering queuing-processes in packet-switched broadcast communication. In *COMPUTER NETWORKS AND ISDN SYSTEMS*, volume 5, pages 58–58. ELSEVIER SCIENCE BV PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS, 1981.
- [60] Ramesh R Rao and Anthony Ephremides. On the stability of interacting queues in a multiple-access system. *Information Theory, IEEE Transactions on*, 34(5):918–930, 1988.
- [61] Boris Solomonovich Tsybakov and Viktor Alexandrovich Mikhailov. Ergodicity of a slotted aloha system. *Problemy Peredachi Informatsii*, 15(4):73–87, 1979.
- [62] Wojciech Szpankowski. Stability conditions for some distributed systems: Buffered random access systems. *Advances in Applied Probability*, pages 498–515, 1994.
- [63] Guy Fayolle, Erol Gelenbe, and Jacques Labetoulle. Stability and optimal control of the packet switching broadcast channel. *Journal of the ACM (JACM)*, 24(3):375–386, 1977.
- [64] Wei Luo and Anthony Ephremides. Stability of n interacting queues in random-access systems. *Information Theory, IEEE Transactions on*, 45(5):1579–1587, 1999.

- [65] TAREK N Saadawi and Anthony Ephremides. Analysis, stability, and optimization of slotted aloha with a finite number of buffered users. *Automatic Control, IEEE Transactions on*, 26(3):680–689, 1981.
- [66] Moshe Sidi and A Segall. Two interfering queues in packet-radio networks. *Communications, IEEE Transactions on*, 31(1):123–129, 1983.
- [67] Hideaki Takagi and Leonard Kleinrock. Mean packet queueing delay in a buffered two-user csma/cd system. *IEEE transactions on communications*, 33:1136–1139, 1985.
- [68] Joy A. Thomas Thomas M. Cover. *Elements of Information Theory*. John Wiley & sons, 2nd edition, 1991.
- [69] Vyacheslav L Girko. A refinement of the central limit theorem for random determinants. *Theory of Probability & Its Applications*, 42(1):121–129, 1998.
- [70] Peter J Smith and Mansoor Shafi. On a gaussian approximation to the capacity of wireless mimo systems. In *Communications, 2002. ICC 2002. IEEE International Conference on*, volume 1, pages 406–410. IEEE, 2002.
- [71] Marco Chiani, Moe Z Win, and Alberto Zanella. On the capacity of spatially correlated mimo rayleigh-fading channels. *Information Theory, IEEE Transactions on*, 49(10):2363–2371, 2003.
- [72] M. R Leadbetter. *Extreme and related properties of random sequences and processes*. Springer-verlag, 1983.
- [73] S.Coles. *An introduction to statistical modeling of extreme values*. Springer-verlag, 2001.
- [74] MR Leadbetter. Extremes and local dependence in stationary sequences. *Probability Theory and Related Fields*, 65(2):291–306, 1983.
- [75] O.Gurewitz J. Kampeas, A. Cohen. Opportunistic Scheduling in Heterogeneous Networks: Distributed Algorithm and System Capacity. 2012.
- [76] MR Leadbetter. Weak convergence of high level exceedances by a stationary sequence. *Probability Theory and Related Fields*, 34(1):11–15, 1976.
- [77] Joseph Kampeas, Asaf Cohen, and Omer Gurewitz. Capacity of distributed opportunistic scheduling in non-homogeneous networks. *Accepted to IEEE Transactions on Information Theory (arXiv preprint arXiv:1202.0119)*, 2014.
- [78] GL O’Brien. Limit theorems for sums of chain-dependent processes. *Journal of applied probability*, pages 582–587, 1974.
- [79] Giuseppe Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *Selected Areas in Communications, IEEE Journal on*, 18(3):535–547, 2000.
- [80] Ury Yechiali and Pinhas Naor. Queuing problems with heterogeneous arrivals and service. *Operations Research*, 19(3):722–734, 1971.
- [81] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Dover Publications, 2012.
- [82] Pavle Mladenović. Extreme values of the sequences of independent random variables with mixed distributions. *Matematički Vesnik*, (212):29–37, 1999.
- [83] James R. Weaver. Centrosymmetric (cross-symmetric) matrices, their basic properties, eigenvalues, and eigenvectors. *The American Mathematical Monthly*, 92(10):pp. 711–717, 1985.
- [84] R. Serfozo. *Basics of Applied Stochastic Processes*. Applied probability. Springer, 2009.

-
- [85] Julian Keilson and Hans Gerber. Some results for discrete unimodality. *Journal of the American Statistical Association*, 66(334):386–389, 1971.